

Reinhard Stockmann | Hubert Ertl (Hrsg.)

Evaluation und Wirkungsforschung in der beruflichen Bildung



Reinhard Stockmann | Hubert Ertl (Hrsg.)

Evaluation und Wirkungsforschung in der beruflichen Bildung

Impressum

Zitiervorschlag:

Stockmann, Reinhard; Ertl, Hubert (Hrsg.): Evaluation und Wirkungsforschung in der beruflichen Bildung. Bonn 2021

1. Auflage 2021

Herausgeber:

Bundesinstitut für Berufsbildung
Robert-Schuman-Platz 3
53175 Bonn
Internet: www.bibb.de

Publikationsmanagement:

Stabsstelle „Publikationen und wissenschaftliche Informationsdienste“
E-Mail: publikationsmanagement@bibb.de
www.bibb.de/veroeffentlichungen

Herstellung und Vertrieb:

Verlag Barbara Budrich
Stauffenbergstraße 7
51379 Leverkusen
Internet: www.budrich.de
E-Mail: info@budrich.de

Lizenzierung:

Der Inhalt dieses Werkes steht unter einer Creative-Commons-Lizenz (Lizentyp: Namensnennung – Keine kommerzielle Nutzung – Keine Bearbeitung – 4.0 International).

Weitere Informationen zu Creative Commons und Open Access finden Sie unter www.bibb.de/oa.

ISBN 978-3-8474-2941-8 (Print)

ISBN 978-3-96208-259-8 (Open Access)

urn:nbn:de:0035-0915-5

Bibliografische Information der Deutschen Nationalbibliothek

Die deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Gedruckt auf PEFC-zertifiziertem Papier

► Inhaltsverzeichnis

Abkürzungsverzeichnis	5
<i>Reinhard Stockmann, Hubert Ertl</i>	
Einleitung: Evaluations- und Wirkungsforschung	7
<i>Reinhard Stockmann</i>	
Systematic Evaluation Analysis als Instrument der Wirkungsevaluation	11
<i>Marc Beutner</i>	
Evaluation in der Berufsbildung: das Dreischalenmodell als Grundlage	25
<i>Wolfgang Meyer</i>	
Wirkungsverläufe messen: eine kleine Sinfonie der Berufsbildungszusammenarbeit	45
<i>Lydia Schulze Heuling, Steffen Wild, Ernst Deuer</i>	
Längere Schulzeit – andere Ansprüche? Wie die Beschuldungsdauer die Prioritäten bei der Arbeitgeberwahl beeinflusst	62
<i>Junmin Li</i>	
Ex-ante-Evaluation als Beitrag zum Policy-Transfer? Eine Analyse am Beispiel der Übertragung des Peer-Review-Verfahrens der Berufsbildung von Deutschland nach China ...	86
<i>Rosemarie Sackmann</i>	
Beziehungen zwischen Ex-ante-Evaluation und formativer Evaluation. Beispiele aus der wissenschaftlichen Begleitung zum Pilotprojekt „Entwicklung und Erprobung eines Konzepts zur Förderung von Anerkennungskosten“	104
<i>Peter Jablonka, Philipp Ulmer</i>	
Evaluation einer beruflichen Weiterbildungsmaßnahme im Kontext der deutsch-portugiesischen Berufsbildungskoperationen	122

Birte Komosin, Henning Kruse

Welchen Beitrag kann ein regionales Monitoring- und Berichtssystem für die
Wirkungsforschung in der Berufsbildung leisten? 141

Andrea Mohorič, Nadja Konrad, Wolfgang Wittig

Evaluation des Förderprogramms „Berufsbildung ohne Grenzen“ zur betrieblichen
Beratung bei der Auslandsmobilität von Auszubildenden und jungen Fachkräften 157

Über die Autorinnen und Autoren 179

Zur Arbeitsgemeinschaft Berufsbildungsforschungsnetz 181

Abstract 185

► Abkürzungsverzeichnis

AG BFN	Arbeitsgemeinschaft Berufsbildungsforschungsnetz
AHK	Außenhandelskammer
ANQEP	Agência Nacional para a Qualificação e o Ensino Profissional (Nationale Agentur für Qualifikation und berufliche Bildung, Portugal)
ATT	Average Treatment Effect on the Treated
BA	Bundesagentur für Arbeit
BAFA	Bundesamt für Wirtschaft und Ausfuhrkontroll
BHO	Bundshaushaltsordnung
BIBB	Bundesinstitut für Berufsbildung
BMBF	Bundesministerium für Bildung und Forschung
BMWi	Bundesministerium für Wirtschaft und Energie
BoG	„Berufsbildung ohne Grenzen“ (Förderprogramm)
DBR	Design-Based-Research
DHBW	Duale Hochschule Baden-Württemberg
DIHK	Deutscher Industrie- und Handelskammertag
ESF	Europäischer Sozialfonds
HEGA	Handlungsempfehlung und Geschäftsanweisung der Bundesagentur für Arbeit
HSS	Hanns-Seidel-Stiftung
IBS	Informations- und Beratungsstelle für Auslandsaufenthalte in der beruflichen Bildung bei der NA beim BIBB
IEFP	Instituto do emprego e formação profissional (Nationales Institut für Beschäftigung und Berufsbildung, Portugal)
KMU	Kleine und mittlere Unternehmen
NA	Nationale Agentur beim BIBB
RCT	Randomized Controlled Trial
SEA	Systematic Evaluation Analysis
SenIAS	Berliner Senatsverwaltung für Integration, Arbeit und Soziales
TAM	Technology-Acceptance-Modell

Reinhard Stockmann, Hubert Ertl

► **Einleitung: Evaluations- und Wirkungsforschung**

Mit diesem Band widmet sich die Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN) der Evaluations- und Wirkungsforschung in der beruflichen Bildung. Fragen zu Methoden der Berufsbildungsforschung ziehen sich wie ein roter Faden durch die Veranstaltungen der AG BFN. Sie wurden in thematisch orientierten Veranstaltungen diskutiert, aber auch in eigens dafür ausgelegten methodisch orientierten Tagungen. So wurde beispielsweise in der AG BFN-Veranstaltung in Wien 2010 die Diskussion um die Polarisierung von Grundlagenforschung und Handlungsforschung in den Fokus gerückt. Dabei bewegte sich die Diskussion von der Polarisierung der beiden Forschungsparadigmen weg und hin zu methodischen Fragen einer anwendungsorientierten Berufsbildungsforschung. Auch 2017 zum 25-jährigen Jubiläum der AG BFN wurde das Thema Berufsbildungsforschung wieder aufgenommen und im Kontext der Multidisziplinarität, Praxisorientiertheit und Evidenzbasierung diskutiert. Erwartungen von Praxis und Politik an die Wissenschaft wurden hierbei erläutert.

Somit knüpft die AG BFN mit dieser Publikation an vorherige Diskussionen zur anwendungsorientierten Berufsbildungsforschung an. Es werden Methoden der Evaluations- und Wirkungsforschung erörtert und aus verschiedenen Blickwinkeln betrachtet. Evaluation, Implementations- und Begleitforschung sowie Wirkungsanalysen sind wissenschaftliche Instrumente, die in der Berufsbildung seit jeher einen hohen Stellenwert einnehmen. Ziel ist es, Ex-ante-Bedingungen für die Entwicklung von Projekten, Programmen, Modellvorhaben oder Systemreformen zu analysieren, den Implementationsprozess wissenschaftlich zu begleiten, die intendierten wie nicht intendierten Wirkungen der Interventionsmaßnahmen zu erfassen und ihren kausalen Ursachenfaktoren zuzuschreiben. Auf diese Weise leistet Wissenschaft einen Erkenntnisbeitrag für die Praxis, die Weiterentwicklung von Projekten und Programmen sowie für politische Strategien.

Einleitend führt Reinhard Stockmann in die Begrifflichkeiten ein und stellt die Gemeinsamkeiten und Unterschiede zwischen Evaluation und Wirkungsforschung heraus. Insbesondere in der Wirkungsforschung gibt es heftige Diskussionen um die besten Methoden zur Erfassung der Wirkungen und zur Identifizierung der Ursache-Wirkungs-Beziehung. Favo-

riert werden häufig experimentelle Verfahren, die durch Kontrollgruppen (*Randomized Controlled Trials* (RCT)) versuchen, Wirkungen und Kausalitäten aufzuzeigen. Diese RCT werden oft als „Königsweg“ oder „Goldstandard“ der Wirkungsforschung bezeichnet. Zwar weisen experimentelle Designs eine hohe interne Validität auf; sie sind aber schwach im Hinblick auf die Generalisierbarkeit von Ergebnissen. Hier kommt die Methode der *Systematic Evaluation Analysis* ins Spiel. In seiner Keynote stellt Reinhard Stockmann das Konzept der **Systematic Evaluation Analysis als Instrument der Wirkungsevaluation** vor, mit dessen Hilfe er vier Erfolgsfaktoren identifiziert, die für die Effektivität, Wirksamkeit und Nachhaltigkeit von Maßnahmen in der Berufsbildung verantwortlich sind.

Marc Beutner stellt in seiner Keynote **Evaluation in der Berufsbildung: das Dreischalenmodell als Grundlage** ein Instrument vor, mit dessen Hilfe eine systematisch-analytische Evaluation im Bereich der Berufsbildung durchgeführt werden kann. Das Dreischalenmodell bietet eine Betrachtung von Maßnahmen und Programmen auf unterschiedlichen Ebenen. Auf der ersten Schale wird entsprechend des zu evaluierenden Gegenstands die Methode festgelegt. Die zweite Schale beinhaltet die berufsbildungsspezifischen Besonderheiten, die als Mindeststandards zu beachten sind. Mit der dritten Schale rücken die Einflussfaktoren und Kontextvariablen in den Vordergrund. Anhand verschiedener Maßnahmen in der Berufsbildung wird gezeigt, dass das Dreischalenmodell ein Orientierungsraster für die Planung, Gestaltung und Ausformung von Evaluationen in der Berufsbildung bietet.

Wolfgang Meyer greift in seinem Beitrag **Wirkungsverläufe messen: eine kleine Sinfonie der Berufsbildungszusammenarbeit** die Frage auf, warum experimentelle Methoden in der praktischen Berufsbildungsarbeit ungeeignet sind, um Wirkungen von Interventionen zu messen. Er argumentiert, dass RCT aufgrund ihrer methodischen Ausrichtung die Realität in der Zusammenarbeit der Berufsbildung nicht abbilden. Anhand eines Beispiels wird ein Ansatz aufgezeigt, der die Erfassung von Wirkungen und ihren Verläufen in der Berufsbildungszusammenarbeit ermöglicht.

Auch Lydia Schulze Heuling, Steffen Wild und Ernst Deuer stellen in ihrem Beitrag **Längere Schulzeit – andere Ansprüche? Wie die Beschulungsdauer die Prioritäten bei der Arbeitgeberwahl beeinflusst** eine Alternative zu den RCT, den Goldstandards, in der Wirkungsforschung vor. Mithilfe eines Matchingverfahrens, dem *Propensity Score Matching*, wird geprüft, wie sich die Beschulungsdauer auf die Beurteilung des Arbeitsplatzes auswirkt.

Berufsbildungskonzepte fokussieren zunehmend nicht nur die Projekt- und Programmebene, sondern auch Systemveränderungen im Rahmen von Reformprozessen. Angesichts niedriger Jugendarbeitslosigkeit und des hohen Kompetenzniveaus der Facharbeiter/-innen in Deutschland stößt das duale Ausbildungssystem auf großes internationales Interesse. Im Rahmen der Entwicklungszusammenarbeit gibt es zahlreiche Versuche, Berufsbildungssysteme umzugestalten. Junmin Li legt in ihrem Beitrag **Ex-ante-Evaluation als Beitrag zum Policy-Transfer? Eine Analyse am Beispiel der Übertragung des Peer-Review-Verfahrens der Berufsbildung von Deutschland nach China** den Fokus auf die Systemevaluation. Sie untersucht, inwiefern die Methode der Ex-ante-Evaluation beim Policy-Transfer eingesetzt

werden kann. Hiermit greift sie eine Methode auf, die im Rahmen von Evaluationsprozessen in der beruflichen Bildung bisher eher wenig genutzt wird.

Im Gegensatz zu den Ex-ante-Evaluationen werden formative Evaluationen häufig genutzt, um Erkenntnisse für den Steuerungsprozess zu gewinnen. Ex-ante-Evaluationen sind nicht nur für die Planung von Vorhaben wichtig, sondern sie können auch als „Baseline“ für Wirkungsstudien dienen. Rosemarie Sackmann zeigt die **Beziehung zwischen Ex-ante- und formativer Evaluation an Beispielen aus der wissenschaftlichen Begleitung zum Pilotprojekt zur Entwicklung und Erprobung eines Konzepts zur Förderung von Anerkennungskosten** auf. Sie beschreibt, wie sich die Ex-ante-Evaluation auf die Gestaltung der prozessbegleitenden formativen Evaluation auswirkt und wie die Ergebnisse der formativen Evaluation zur Korrektur der Ex-ante-Evaluation oder als Vorlage für die Ex-ante-Evaluation von Nachfolgeprojekten dienen können.

Während in den vorherigen Beiträgen die Reflexion der Methoden der Evaluation und Wirkungsforschung im Fokus stand, rückt in den letzten drei Beiträgen die Anwendungspraxis in den Vordergrund. Peter Jablonka und Philipp Ulmer stellen in ihrem Beitrag die **Evaluation einer beruflichen Weiterbildungsmaßnahme im Kontext der deutsch-portugiesischen Berufsbildungs Kooperationen** vor. Sie gehen der Frage nach, wie der Transfer von einem Land in ein anderes als gelungen angesehen werden kann. Hierbei greifen sie auf das von Reinhard Stockmann in diesem Band vorgestellte Konzept der Schlüsselgrößen für Nachhaltigkeit zurück.

Birte Komosin und Henning Kruse stellen sich die Frage, **welchen Beitrag ein regionales Monitoring- und Berichtssystem für die Wirkungsforschung in der Berufsbildung leisten kann**. Im Vordergrund stehen öffentlich geförderte Beratungs- und Coachingleistungen in der beruflichen Weiterbildung, deren Zielerreichungen durch ein digitales Monitoring- und Berichtssystem sichtbar gemacht werden. Erörtert wird, wie ein solches System für weitergehende Evaluierungen und Wirkungsuntersuchungen genutzt werden kann.

Auch Andrea Mohorič, Nadja Konrad und Wolfgang Wittig berichten direkt aus der Praxis. Sie stellen die **Evaluation des Förderprogramms „Berufsbildung ohne Grenzen“ zur betrieblichen Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften** vor. Die Evaluation orientiert sich an den Vorgaben der Bundeshaushaltsordnung. Es wurden Handlungsempfehlungen für die Fortsetzung des Förderprogramms aufgrund der Evaluationsergebnisse aufgestellt.

Die vorliegenden Beiträge geben somit Einblicke in unterschiedliche Ansätze der Evaluation und Wirkungsforschung. Unser Dank gilt den Gutachterinnen und Gutachtern für die hilfreichen Kommentare, den Autorinnen und Autoren für die gute Zusammenarbeit, Frau Dr. Widera vom BIBB für ihre umsichtige Begleitung aller das Forum und diesen Band betreffenden Prozesse sowie allen Teilnehmenden des AG BFN-Forums für die anregenden Diskussionen.

Reinhard Stockmann

► Systematic Evaluation Analysis als Instrument der Wirkungsevaluation

In diesem Beitrag werden die konzeptionellen Gemeinsamkeiten und Unterschiede zwischen Evaluation und Wirkungsforschung sowie ihrer methodologischen Ansätze bei der Kausalanalyse erörtert. Um aus Evaluationen und Wirkungsstudien systematisch Wissen zu kumulieren, gibt es unterschiedliche Ansätze, z. B. Metaevaluation, Metaanalyse, *Literature Review*, *Systematic Review* und *Research Synthesis*. Diese Konzepte sind nicht trennscharf und werden in unterschiedlicher Weise definiert.

Nach einer kurzen Begriffsklärung wird eine Analysemethode vorgestellt, die Elemente dieser verschiedenen Konzepte miteinander vereint und deshalb als „*Systematic Evaluation Analysis*“ bezeichnet wird. Diese Methodik wird anschließend dazu verwendet, eine Reihe von seit 1992 selbst durchgeführten Evaluationen zur deutschen Berufsbildungszusammenarbeit zu analysieren, um die Erfolgsfaktoren der Zusammenarbeit zu identifizieren. Diese werden zu vier Schlüsselgrößen geclustert: (1) Systemkompatibilität, (2) Flexible Steuerung, (3) Personal und (4) *Ownership*. Aus diesen wird ein Modell für eine nachhaltige Berufsbildungszusammenarbeit entwickelt.

1 Evaluation und Wirkungsforschung

Bevor ich zu meinem eigentlichen Thema komme, lassen Sie mich zuerst etwas zu dem Tagungsthema: „Evaluation und Wirkungsforschung in der Berufsbildung“ sagen. Zunächst einmal ist festzuhalten, dass es sich bei der Evaluation um das breitere Wissenschaftskonzept handelt. Während die Wirkungsforschung eine klassische Disziplin ist, die seit alters die Zusammenhänge von Ursache und Wirkung untersucht, ist die Evaluation ein Kind der Moderne (vgl. STOCKMANN 2007; STOCKMANN/MEYER 2014; STOCKMANN/MEYER 2017). Diese zeichnet sich dadurch aus, dass traditionell und religiös bestimmte Glaubensvorstellungen von der Ordnung der Welt durch das Vertrauen auf Rationalität und Fortschritt ersetzt werden (vgl. LERNER 1968, S. 387). Damit verbunden ist die Überzeugung, dass die Moderne gestaltet und gesteuert werden kann. Sollen soziale Veränderungsprozesse eingeleitet werden, um gesellschaftlich akzeptierte Ziele zu erreichen, z. B. Chancengleichheit, soziale Gerechtigkeit, nachhaltige Entwicklung etc., sind Informationen (Daten, Fakten, *Evidences*) notwendig, um solche Prozesse steuern zu können. Es gilt zu erkennen, welche Faktoren

welche Wirkungen hervorrufen, damit Interventionen unter gegebenen Rahmenbedingungen gezielt ausgewählt und implementiert werden können.

Hier kommen nun die Evaluation und die Wirkungsforschung ins Spiel. Beide sind geeignet – neben anderen Instrumenten – *Evidences* zu produzieren, mit denen nicht nur die Effektivität – also die Zielerreichung –, sondern auch die Wirksamkeit von *Policies*, Strategien, Programmen, Projekten und einzelnen Maßnahmen bestimmt und auf ihre kausalen Ursachenfaktoren zurückgeführt werden kann. Darüber hinaus können mithilfe von Evaluationen Kriterien wie Relevanz, Effizienz, Kohärenz, Signifikanz, Nachhaltigkeit etc. untersucht und bewertet werden.

Evaluation ist anders als die Wirkungsforschung nicht nur retrospektiv aufgestellt, sondern wird in Form von Ex-ante-Evaluationen auch planerisch tätig. Mit Ex-ante-Evaluationen werden Informationen und Bewertungen für die Planungsphase, für die aktive und prozessorientierte Gestaltung von Strategien und Prozessen geliefert. Es handelt sich um „*Analysis for policy*“ oder „*Science for action*“. In der Implementierungsphase kann diese formative Prozessperspektive beibehalten und zusätzlich um eine summative Komponente erweitert werden, indem die Durchführungsprozesse und ihre Resultate beobachtet werden (*On-going-Evaluation*). Des Weiteren – und da trifft sich die Evaluationsforschung wieder mit der Wirkungsforschung – kann Evaluation dazu dienen, summativ, bilanzierend und ergebnisorientiert Wirkungen zu beobachten und ihren Ursachenfaktoren zuzuschreiben. In diesem Sinne betreibt Evaluation dann „*Analysis of policy*“ oder „*Science for knowledge*“. Werden solche Evaluationen ex post durchgeführt, lässt sich auch die Nachhaltigkeit von Interventionen messen und bewerten.

Tabelle 1: Gemeinsamkeiten und Unterschiede zwischen Evaluation und Wirkungsforschung

	Evaluation	Wirkungsforschung
Herkunft	neue Disziplin	alte Disziplin
Spektrum	Planung, Implementation, Wirkung formativ + summativ	Wirkung summativ
Erkenntnisinteresse	<i>Science for action</i> + <i>Science for knowledge</i>	<i>Science for knowledge</i>
Ziel	Nutzen für Entscheider/-innen stiften, Bewertungen und Empfehlungen geben	Kausalität offenlegen
Methodik	empirisch wissenschaftliche Forschungslogik + Handlungsforschung	empirisch wissenschaftliche Forschungslogik

Evaluation beinhaltet demnach immer eine Analyse- und eine Bewertungskomponente. Für die Analyse bedient sie sich – wie die Wirkungsforschung – aller Instrumente der empirischen Sozialforschung und ist auch deren wissenschaftlichen Qualitätsstandards wie Objektivität, Reliabilität und Validität verpflichtet. Doch über diese wissenschaftliche Wirkungsanalyse hinaus ist jede Evaluation mit einer Bewertung verbunden. Jede gängige Evaluationsdefini-

tion enthält deshalb im Kern diesen Aspekt. So z. B. in der vielzitierten Definition von Donna Mertens (1998, S. 219): „Evaluation is the systematic investigation of the merit or worth of an object (program)“, zumeist zu dem Zweck: „of reducing uncertainty in decision making“.

Für die Bewertung des Wertes oder Nutzens eines Evaluationsobjektes werden Kriterien benötigt. Diese können, im Unterschied zu Normenreihen wie ISO, ganz verschieden sein. Einige wie die Relevanz, Effektivität, Effizienz, Wirksamkeit, Nachhaltigkeit von Projekten, Programmen etc. habe ich schon genannt. Wichtig ist, dass die Bewertungskriterien offen gelegt und klar definiert werden, sodass sie intersubjektiv nachvollzogen werden können. Wichtig ist weiterhin, dass Evaluation Bewertungen (z. B. anhand solcher Kriterien) vornimmt, aber keine Werturteile fällt. Ansonsten wäre jede Bewertung subjektiv und von dem Wertesystem des oder der Bewertenden abhängig.

Zusammenfassend kann festgehalten werden, dass Evaluation im Vergleich zur Wirkungsforschung das breitere Wissenschaftskonzept ist, das nicht nur summativ Wirkungen erfasst und sie ihren Kausalfaktoren zuschreibt, sondern sich auch auf andere Programmphasen oder Aspekte beziehen kann. Zudem kommt Wirkungsforschung auch ohne die Bewertung des Nutzens der untersuchten Interventionen aus. Sie kann sich allein auf die Wirkungsbeobachtung und Kausalanalyse beschränken. Anders als die Evaluation kann Wirkungsforschung auch als Grundlagenwissenschaft betrieben werden, also ohne das Ziel, in einem Programm oder Projekt einen konkreten Nutzen zu stiften, z. B. um wissenschaftlich gesicherte Erkenntnisse für Entscheidungsprozesse zu liefern.

Schaut man sich an, wie die beiden Konzepte in der Berufsbildungsforschung verwendet werden, dann fällt auf, dass Wirkungsforschung häufig dem Kontext der empirisch analytischen Grundlagenforschung und Evaluation eher dem Konzept der zumeist qualitativ ausgerichteten Handlungsforschung zugeschrieben wird (vgl. SEVERING/WEISS 2012, S. 6; EULER 2018, S. 42f., 45). Diese Typisierung wird jedoch der Evaluationsforschung nicht gerecht, die zwar auch das Konzept der Handlungsforschung nutzt, sich aber grundsätzlich den Prinzipien der empirisch wissenschaftlichen Sozialforschung verpflichtet sieht, wie man z. B. den Standards für Evaluation der DeGEval (2016) entnehmen kann.

2 Kausalproblematik in der Wirkungsforschung

Wenn es um Wirkungsforschung oder Wirkungsevaluation geht, muss zunächst einmal geklärt werden, was die Wirkung einer Intervention darstellt. Häufig werden Wirkungen mit Outputs verwechselt. Outputs sind die Leistungen, die als Folge von Programmaktivitäten erbracht werden, also z. B. die Zahl Kranker, die in einer Gesundheitsstation behandelt werden, oder die Zahl durchgeführter Umweltberatungen einer Handwerkskammer.

Was uns aber in der Wirkungsevaluation oder Wirkungsforschung interessiert, sind die daraus entstandenen Wirkungen, also z. B. die Verbesserung des Gesundheitszustands der behandelten Kranken oder das veränderte Umweltverhalten der Beratenen. Dabei steht die Wirkungsevaluation vor zwei Herausforderungen: zum einen der Erfassung möglichst aller

messbaren Wirkungen, d. h. sowohl der intendierten als auch der nicht intendierten, sowie zweitens der Identifizierung der Ursachenfaktoren. Bei der Kausalanalyse geht es vor allem darum, alternative Erklärungen auszuschließen. So könnte ein krisenhafter Zustand von alleine wieder verschwinden. In der Medizin nennt man das „spontane Remission“. Diese wäre z. B. dann gegeben, wenn nicht die Behandlung des Arztes, sondern die Selbstheilungskräfte des Körpers zu einer Verbesserung des Gesundheitszustands geführt haben. Darüber hinaus können auch exogene Ereignisse für beobachtete Veränderungen verantwortlich sein. So könnten z. B. nicht die Umweltberatungen das beobachtete veränderte Verhalten ausgelöst haben, sondern gesetzliche Vorgaben oder allgemeine strukturelle Trends.

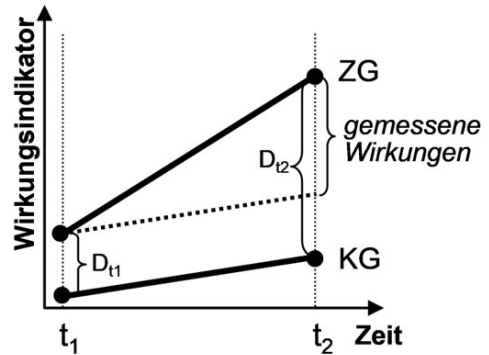
Fassen wir zusammen: Ziel der Wirkungsevaluationen ist einerseits, möglichst viele, im Idealfall alle intendierten und nicht intendierten Wirkungen zu erfassen. Hierfür ist der Einsatz von möglichst vielen unterschiedlichen Datenerhebungsmethoden, wie z. B. standardisierte Interviews, leitfadengestützte Interviews, Fokusgruppen, Beobachtungen, Sekundärdaten- oder Textanalysen etc., notwendig. Andererseits geht es bei der Wirkungsevaluation darum, konkurrierende Erklärungen auszuschließen, um Ursache-Wirkungs-Zusammenhänge optimal zu belegen. Für diesen Zweck bietet sich eine Reihe von Untersuchungsdesigns an; auf alle kann ich hier nicht eingehen. Wegen ihrer hohen internen Validität werden experimentelle Verfahren bevorzugt, am besten als *Randomized Controlled Trials*, bei denen man sich sicher sein kann, dass die beobachtete Wirkung auf eine bestimmte Intervention zurückgeführt werden kann. Allerdings ist die externe Validität von Experimenten nur schwach ausgeprägt, da sich alles darum dreht, in einer künstlichen, laborartigen Situation alle Rahmenbedingungen konstant zu halten, sodass sich Versuchs- und Kontroll-(Vergleichs-)Gruppe nur in einem Punkt unterscheiden: der Intervention. In der realen Welt lassen sich jedoch solche Rahmenbedingungen häufig nicht herstellen, sodass die externe Validität, also die Sicherheit, dass der Kausalzusammenhang, der zwischen zwei Variablen festgestellt wird, auch in anderen Rahmenbedingungen gültig ist, eher gering ist.

Grundsätzlich geht es bei allen Kausaluntersuchungen immer darum, alternative Erklärungen auszuschließen und die Frage zu beantworten, was gewesen wäre, wenn es keine Intervention gegeben hätte. Eine solche kontrafaktische Situation ist gegeben, wenn es gelingt, eine Gruppe (Kontroll- oder Vergleichsgruppe) zu bilden, die so gut wie irgend möglich der Versuchsgruppe ähnelt, das heißt am besten mit dieser in allen wichtigen Charakteristika übereinstimmt. Wenn wir feststellen, dass sich Veränderungen nur bei der Versuchsgruppe einstellen, nicht aber bei der Kontrollgruppe, dann kann man mit großer Sicherheit sagen, dass der Unterschied auf die Intervention zurückzuführen ist. Um sicherzustellen, dass Kontroll- und Versuchsgruppe sich nicht unterscheiden, wird die Technik der Randomisierung angewendet. Ist diese nicht anwendbar, können Kontrollgruppen auch anders gebildet werden, dann spricht man von Vergleichsgruppen und einem quasiexperimentellen Design. Ungeachtet dessen wird ein sogenannter *Double-Difference*-Ansatz eingesetzt, bei dem aktuell erhobene Daten (t_2) mit Baseline-Daten (vor der Intervention) (t_1) bei der Versuchs- und Kontroll- bzw. Vergleichsgruppe miteinander verglichen werden, um Veränderungen feststellen zu können.

Abbildung 1: Experimenteller Ansatz

Double-Difference-Ansatz (t_1 und t_2)
Vergleich aktuell erhobener Daten (t_2)
mit Baseline- & Monitoring-Daten (t_1)
bei der Ziel- und Vergleichsgruppe zur
Bestimmung des tatsächlichen
Ausmaßes der Wirkungszuschreibung

Zeitpunkt (t_1) z. B. Baseline
(t_2) z. B. Survey-Daten



ZG: Zielgruppe, KG: Kontrollgruppe,
D: Differenz ZG – KG, t: Zeitpunkt

Quelle: CASPARI/BARBU 2008, S. 9

Nun gibt es in der Realität leider häufig das Problem, dass kein randomisiertes Experiment oder auch kein Quasiexperiment mit Vergleichsgruppen eingesetzt werden kann, z. B. wenn

- ▶ eine Intervention sich auf ein ganzes Land bezieht, weil von einem Gesetz, einer Regelung oder einer Versicherungsleistung alle betroffen sind, oder
- ▶ wenn aus ethischen Gründen keine Personen zu Versuchszwecken ausgeschlossen werden können (z. B. im Kontext humanitärer Hilfe oder Medikamentenversorgung) oder
- ▶ wenn sich keine Gruppen finden lassen, die nicht in irgendeiner Form gefördert wurden (z. B. in Ländern mit einer sehr hohen Geburtenrate), oder
- ▶ wenn Baseline-Daten nicht erhoben wurden und sich auch nachträglich nicht mehr valide retrospektiv erheben lassen.

In solchen, nicht seltenen Fällen müssen andere Untersuchungsdesigns für die Kausalanalyse herangezogen werden, z. B. ein *Single-Difference*-Ansatz wie die Querschnittsanalyse, bei dem nur ein Vergleich von Versuchs- und Kontrollgruppe zu einem Zeitpunkt (t_2) stattfindet, oder ein *Pretest-Posttest*-Design, bei dem nur ein Vorher-nachher-Vergleich bei der Versuchsgruppe stattfindet, oder ein um weitere Messzeitpunkte erweitertes Paneldesign, eine Zeitreihenanalyse oder eben ein Verfahren, das in diesem Beitrag vorgestellt werden soll: die *Systematic Evaluation Analysis*.

3 Metaevaluation, Systematic Review und Systematic Evaluation Analysis

Zuerst ist eine kleine Begriffsexplikation vonnöten, da die folgenden Begriffe in einer verwirrenden, teils sich widersprechenden Vielfalt verwendet werden. Dennoch haben Metaevaluation, Metaanalyse, *Literature Review*, *Systematic Review* und *Research Synthesis* ein Ziel gemeinsam: Sie wollen Wissen systematisch kumulieren.

Nachdem der Begriff **Metaevaluation** zunächst sowohl für die Kumulation inhaltlicher Befunde aus Evaluationsstudien als auch für die Bewertung ihrer methodischen Qualität verwendet wurde, hat sich mittlerweile eine terminologische Differenzierung durchgesetzt (vgl. STOCKMANN/SILVESTRINI 2013). Demnach wird der Begriff Metaevaluation nur noch für die Evaluation von Evaluationen¹ (vgl. SCRIVEN 1991, S. 228ff.; WIDMER 1996, S. 4; CASPARI 2009, S. 2f.) als eine Form der Qualitätsprüfung von Evaluation verwendet. Diese erfolgt häufig entlang der Standards für Evaluationen.

Der Begriff **Metaanalyse** bezeichnet hingegen eine Vorgehensweise, die systematisch die Ergebnisse verschiedener Einzelstudien aus einem Forschungsbereich integriert und subsumiert, also rein inhaltlich ausgerichtet ist.

Diesem Ansatz ähnelt das **Systematic Review**, das das Ziel verfolgt, „to locate, appraise and synthesize the best available evidence relating to a specific research question in order to provide informative and evidence-based answers“ (BOLAND/CHERRY/DICKSON 2017, S. 2). *Systematic Reviews* „typically involve a detailed and comprehensive plan and search strategy derived a priori, with the goal of reducing bias by identifying, appraising, and synthesizing all relevant studies on a particular topic“ (UMAN 2011, S. 57).

Mitunter werden *Systematic Reviews* als Untertyp des **Literature Review** bezeichnet (COOPER 2017, S. 1ff.), die bereits publiziertes Material auswerten: „not based primarily on new facts and findings but on publications containing such primary information, whereby the latter is digested, sifted, classified, simplified, and synthesized“ (MANTEN 1973, S. 75).

Von diesen Begriffen wird wiederum der Begriff **Research Synthesis** mit dem Verweis auf einen anderen Fokus und ein anderes Ziel abgegrenzt:

„Research syntheses attempt to integrate empirical research for the purpose of creating generalizations. [...] Also, research synthesis almost always pay attention to relevant theories, critically analyze the research they cover, try to resolve conflicts in the literature, and attempt to identify central issues for future research.“ (COOPER/HEDGES 2009, S. 6)

Hier wird eine Analysemethode angewendet, die sowohl Elemente der Metaevaluation als auch der Metaanalyse bzw. der damit verwandten Review-Konzepte umfasst und deshalb als **Systematic Evaluation Analysis (SEA)** bezeichnet werden soll. Mit der Metaevaluation hat SEA gemeinsam, dass es sich um die Analyse von Evaluationen handelt, deren metho-

1 Schon 1940 hat Orata seinen Aufsatz mit „Evaluating Evaluation“ (ORATA 1940, S. 641) für diesen Sachverhalt überschrieben.

dische Qualität geprüft wird, um dann inhaltliche Ergebnisse nach ausgewählten Kriterien oder Hypothesen zu synthetisieren, zu strukturieren, zu klassifizieren und zu subsumieren, um letztlich kumulierte, verdichtete Befunde zu generieren, die Generalisierungen zulassen. Eine Evaluation der methodischen Qualität der dafür herangezogenen Studien ist notwendig, da eine inhaltliche Kumulation von Ergebnissen nur auf der Basis methodisch einwandfrei durchgeführter Studien, die eine hohe Reliabilität und Validität aufweisen, sinnvoll ist.

Bei der hier vorgenommenen Analyse handelt es sich nicht um ein *Literature Review*, sondern um die Synthetisierung einer Reihe von selbst durchgeführten Evaluationen zur deutschen Berufsbildungszusammenarbeit (vgl. Tabelle 2).

Tabelle 2: Empirische Grundlage: Analyse bilateraler Berufsbildungskoperationen

Jahr	Autoren/Autorinnen	Analysierte bilaterale Berufsbildungsmaßnahmen
1992	Stockmann	9 in Lateinamerika
1996	Stockmann	15 in Lateinamerika
1997	Stockmann, Leicht	1 in Afrika
1998	Stockmann, Kohlmann	6 in Afrika 6 in Lateinamerika 5 in Asien (davon eine in China)
2000	Stockmann, Meyer, Krapp, Köhne	8 in Asien (alle in China)
2013	Stockmann, Silvestrini	6 in Asien (davon eine in China) 3 in Afrika 2 in Europa 1 in Lateinamerika
2017	Stockmann, Meyer	21 in Asien (alle in China)

Ziel ist es, die Erfolgsfaktoren zu identifizieren, die für die Effektivität, Wirksamkeit und Nachhaltigkeit von Berufsbildungsprojekten verantwortlich sind. Während es in den ausgewählten Studien selbst vor allem darum ging, die interne Validität der Aussagen zu stärken, und dementsprechend – wenn möglich – quasiexperimentelle Untersuchungsansätze verwendet wurden, geht es jetzt darum, ein Verfahren anzuwenden, das eine hohe externe Validität aufweist, um zu generalisierbaren Aussagen zu kommen. Hierfür wird so vorgegangen, dass systematisch untersucht wird, welche Faktoren für den „Erfolg“ eines Projekts oder Programmes zwingend erforderlich sind. Hierfür werden die in den einzelnen Evaluationen identifizierten Erfolgsfaktoren daraufhin abgeglichen, wie häufig sie aufgetreten sind. Es wird versucht herauszufinden, ob es sich nur um individuelle Erfolgsfaktoren handelt, die nur einmal im spezifischen Programm- oder Länderkontext auftreten, oder ob es sich um Faktoren handelt, die stets auftreten, wenn es sich um ein erfolgreiches Projekt handelt. In diesem Fall wurde gefolgert, dass es sich unabhängig vom jeweiligen Programmtypus und unabhängig vom jeweiligen Programmkontext um einen generalisierbaren, d. h. zwingend notwendigen Erfolgsfaktor handelt.

Für die Analyse wurden sieben in den Jahren 1992 bis 2017 durchgeführte Studien herangezogen, in denen 83 Projekte und Programme der Berufsbildungszusammenarbeit weltweit mit dem gleichen Konzept, dem CEval-Evaluationsansatz, untersucht wurden.

Da die Berufsbildungsförderung zum Urgestein deutscher Entwicklungspolitik gehört, also schon über eine 50-jährige Historie verfügt, ist es nicht verwunderlich, dass die Förderziele und die Maßnahmen, um diese zu erreichen, in den vergangenen Dekaden einem deutlichen Wandel unterlagen. Waren die 1960er- und 1970er-Jahre zunächst noch durch die Förderung vollschulischer Ausbildungsmodelle geprägt, wurden in den 1980er- und 1990er-Jahren Systemreformen mit dem Ziel, duale oder kooperative Ausbildungsformen einzuführen, unterstützt. Spätestens um die Jahrtausendwende wurde die Berufsbildung stärker in den Fokus von Wirtschafts- und Beschäftigungsförderung gestellt und am Leitbild einer sozialen und ökologischen Marktwirtschaft ausgerichtet. Zusätzlich sollte die Berufsbildung aber auch direkt zur Armutsbekämpfung beitragen.

Klassifiziert man die untersuchten Projekte, lassen sich drei unterschiedlich starke Gruppen ausmachen. Am umfangreichsten sind die Institutionenförderprojekte, die ihre Blütezeit zwischen 1960 und 1980 hatten und die danach von systemreformerischen Projekten abgelöst wurden, ohne allerdings ihre wichtige Bedeutung bis heute zu verlieren. Projekte für marginalisierte Gruppen kamen vermehrt erst in neuerer Zeit auf und spielen insgesamt im Berufsbildungsportfolio nach wie vor eine marginale Rolle.

Aufzählung der untersuchten Projekte:

- ▶ Institutionenförderprojekte
 - ▶ Singuläre Projekte wurden unterstützt, um deren Leistungsangebot zu verbessern
- ▶ „Kooperative“ Ausbildungsprojekte
 - ▶ Erhöhung des Praxisanteils in der beruflichen Ausbildung und Verbreitung des Ausbildungstyps durch Systemreformen
- ▶ Marginalisierte Gruppen
 - ▶ Projekte im Zusammenhang mit friedensschaffenden, konfliktreduzierenden Maßnahmen
 - ▶ Erwachsenenbildung

Insgesamt lässt sich feststellen, dass die Ansprüche an Berufsbildungsvorhaben im Zeitverlauf immer höhergeschraubt wurden. So sollen sie einerseits dazu beitragen, in den Partnerländern soziale und ökologische Marktwirtschaften aufzubauen, andererseits aber auch dabei helfen, Armut zu reduzieren. Diese Ziele sind nur schwer miteinander kompatibel, da im ersten Fall vor allem systemreformerische Ansätze auf makrogesellschaftlicher Ebene gefragt sind, für die zweite Zielvorgabe jedoch vor allem Projekte im informellen Sektor und auf mikrogesellschaftlicher Ebene vonnöten sind. Die im Zeitverlauf zu beobachtende zunehmende Komplexität von Programmen und Projekten hin zu einem vielfältigen, ausdiffe-

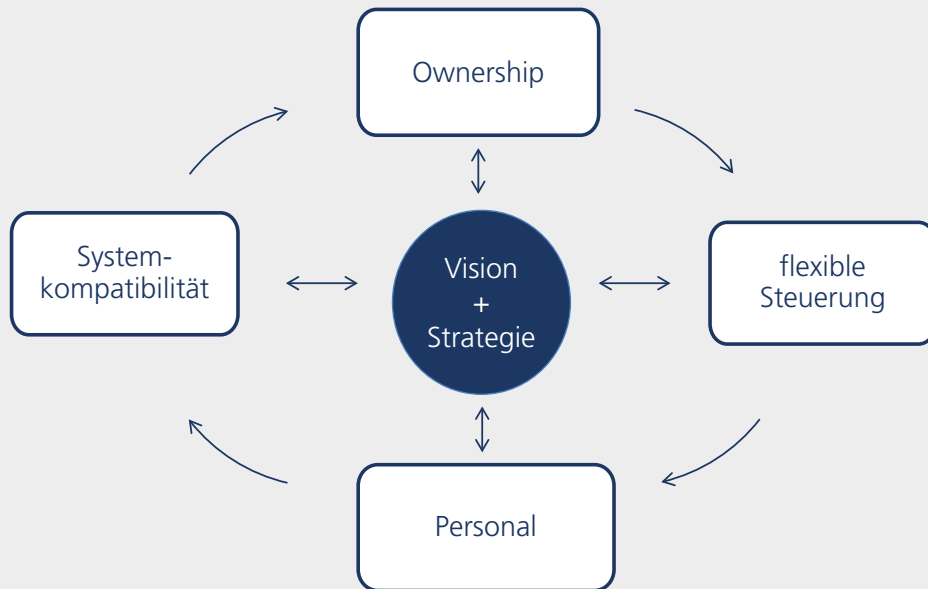
renzierten Zieletableau hat sich nicht unbedingt ausgezahlt: Die Evaluationsstudien zeigen, dass weniger ambitionierte Projekte, z. B. die Förderung ausgewählter Berufsbildungseinrichtungen (Institutionen-Förderprojekte) ohne Breitenwirksamkeits- und Systemveränderungsanspruch, und Projekte, die ohne einen komplexen Mehrebenenansatz implementiert werden konnten, im Hinblick auf das Nachhaltigkeitskriterium am besten abschneiden. Dieser Befund sollte natürlich nicht dahingehend interpretiert werden, möglichst „einfache“ Projekte und Programme durchzuführen, da diese am erfolgversprechendsten sind, sondern komplexe Projekte, also z. B. solche mit einem Systemveränderungsanspruch, vor Beginn besonders genau auf ihre Erfolgchancen, z. B. mithilfe einer Ex-ante-Evaluation, zu prüfen.

Dabei sollten die in unserer *Systematic Evaluation Analysis* herausgearbeiteten Erfolgsfaktoren beachtet werden. Hierfür wurde die Vielfalt der in den einzelnen Projekt- und Programmevaluationen ermittelten Erfolgsgrößen in einem, wie eingangs dargestellt, vergleichenden Verfahren auf vier Kerngrößen verdichtet² (vgl. Abbildung 2). Diese vier „Schlüsselgrößen“ sind:

1. **Systemkompatibilität:** Die im Rahmen der Berufsbildungszusammenarbeit eingeführten Maßnahmen, Institutionen und Organisationen müssen mit dem bestehenden Berufsbildungssystem vereinbar sein oder dieses entsprechend einer besseren Passung verändern.
2. **Flexible Steuerung:** Die Vorhaben der Berufsbildungszusammenarbeit müssen auf veränderte Rahmenbedingungen, technische Innovationen, wirtschaftliche Krisen, neue Kompetenzanforderungen und/oder veränderte Bedürfnisse der Beteiligten schnell und angemessen reagieren und die implementierten Maßnahmen, Institutionen und Organisationen entsprechend den veränderten Anforderungen anpassen.
3. **Personal:** Die mit der Durchführung betrauten organisatorischen Einheiten müssen entsprechend den Anforderungen mit angemessenen Ressourcen ausgestattet sein, wovon in allererster Linie geeignetes Personal mit den benötigten Kompetenzen zu verstehen ist.
4. **Ownership:** Die Trägerorganisationen von Programmen und Projekten müssen sich die Vorhaben zu eigen machen und diese nicht als „von außen“ (insbesondere von den ausländischen Kooperationspartnern) aufgesetzt begreifen, damit sie dauerhaft – auch nach Ende der Vorhabenlaufzeit – mit den notwendigen Ressourcen ausgestattet sind und die zur Erzielung von Wirkungen erforderlichen Entscheidungen getroffen werden.

2 Dieses Vorgehen orientiert sich am Modell der *Balanced Scorecard* („ausgewogener Berichtsbogen“). Hierbei handelt es sich um ein Konzept zur Messung, Dokumentation und Steuerung von Unternehmensaktivitäten. Es dient als Führungsinstrument zur Steuerung von Organisationen anhand strategischer Ziele (vgl. KAPLAN/NORTON 1997; 2001; FRIEDAG/SCHMIDT 2011; GLEICH 2012).

Abbildung 2: Kernschlüsselgrößen



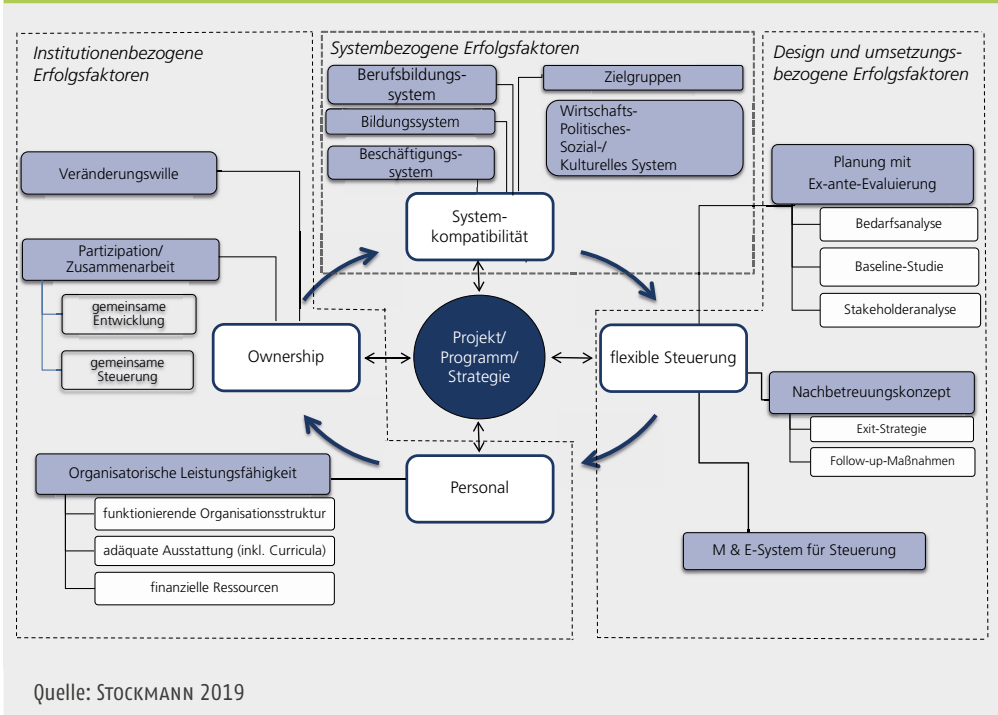
Quelle: STOCKMANN 2019

Diese Kerngrößen werden von drei unterschiedlichen Bündeln (Clustern) von Einflussfaktoren geprägt, wobei die Effektstärke zwischen unterschiedlichen Vorhabenarten variieren kann. Diese Cluster sind (vgl. Abbildung 3):

1. **Systembezogene Erfolgsfaktoren:** Dies sind alle Einflüsse, die aus dem Berufsbildungssystem sowie den mit ihm verbundenen anderen gesellschaftlichen Systemen (speziell dem Beschäftigungs-, dem Bildungs-, dem politischen und dem Wirtschaftssystem) resultieren. Hier handelt es sich z. B. um kulturelle Eigenarten, gesetzliche Vorgaben, Curricula, gesellschaftliche Regeln, Verwaltungsmaßnahmen, wirtschaftliche und politische Entwicklungen.
2. **Design- und umsetzungsbezogene Einflussfaktoren:** Dies sind alle Einflüsse, die sich aus der Planung, Konzeption, Durchführung und Steuerung der Berufsbildungsvorhaben ergeben. Hier handelt es sich z. B. um die bereitgestellten Ressourcen, die Laufzeit, die implementierten Entscheidungsstrukturen, das Monitoring- und Evaluationssystem, Führungsstil und Führungskultur innerhalb der implementierten Strukturen.
3. **Institutionenbezogene Erfolgsfaktoren:** Dies sind alle Einflüsse, die aus den Strukturen, Prozessen, Regeln und Verhaltensweisen der beteiligten Organisationen und Netzwerken abzuleiten sind. Unter diesen Faktoren sind die Leistungsfähigkeit der Trä-

gerorganisationen, die formale und informelle Kooperation in der bilateralen Zusammenarbeit zwischen allen beteiligten Stakeholdern sowie deren Einstellungen gegenüber dem Vorhaben und den implementierten Maßnahmen zu verstehen.

Abbildung 3: Modell der Schlüsselgrößen nachhaltiger Berufsbildungszusammenarbeit



Aus diesem sehr allgemeinen und auf einer Vielzahl unterschiedlicher Vorhaben der Berufsbildungszusammenarbeit beruhenden Schema lassen sich Thesen und Erwartungen über die Wirkungen und die fördernden bzw. behindernden Faktoren konkreter Einzelbeispiele ableiten. An dieser Stelle seien als Beispiele nur einige zentrale Thesen des Modells angesprochen:

- Von den design- und umsetzungsbezogenen Faktoren ist eine flexible Steuerung am wichtigsten, die auf neue Informationen und sich verändernde Umweltbedingungen schnell und angemessen reagiert. Die hierfür benötigten Gestaltungsspielräume und schlanken Entscheidungsstrukturen sind für den Erfolg wichtiger als eine sorgfältige, umfassende und detaillierte Planung. Planungsfehler können durch eine flexible Steuerung eher aufgefangen werden als umgekehrt gut geplante, aber an den realen Erfordernissen vorbeigehende Maßnahmen, die entsprechend dem Plan akribisch abgearbeitet werden. In der

Durchführungsphase ist deshalb vor allem eine rechtzeitige Bereitstellung relevanter und korrekter Monitoring- und Evaluationsergebnisse für die notwendigen Entscheidungen wichtig, durch die Steuerungsfehler vermieden oder zumindest begrenzt werden können. Bei Auslaufen von Fördermitteln und einem entsprechenden Übergang ist die rechtzeitige Implementierung angemessener Überleitungsmaßnahmen entscheidend.

- ▶ Der zentrale Erfolgsfaktor jeder Berufsbildungsmaßnahme ist deren Systemkompatibilität, also eine hohe Vereinbarkeit mit bzw. eine Anschlussfähigkeit an bestehende Strukturen und Abläufe. Als kritischer Aspekt bei der Etablierung einer praxisnahen beruflichen Qualifizierung hat sich vor allem die Balance im Verhältnis Arbeitgeber, Arbeitnehmer und Staat bezüglich der Aus- und Weiterbildungsverantwortung herausgestellt. Hier helfen gut etablierte Vertretungsstrukturen und Institutionen zur Regulierung des Interessenausgleichs weiter, die allerdings eine entsprechende Vertretungs- und Verhandlungskultur voraussetzen. Dies lässt sich nicht einfach durch staatliche Maßnahmen von oben „verordnen“, sondern erfordert einen langwierigen und in der Regel schwierigen Einstellungs- und Kulturwandel, welcher durch entsprechende, den Diskurs fördernde Maßnahmen unterstützt werden muss.
- ▶ Die Umsetzung von Maßnahmen innerhalb eines Vorhabens der Berufsbildungszusammenarbeit erfordert schließlich den Aufbau eines entsprechenden institutionellen Rahmens, der vor allem durch das gegenseitige Vertrauen in die Zusammenarbeit sowie die Bereitschaft des Partners geprägt ist, die im Rahmen der Kooperation entwickelten Maßnahmen und Institutionen eigenständig auch nach Auslaufen der Förderung fortzuführen. Häufig bleiben Projekte und Programme „Fremdkörper“ innerhalb der Trägerorganisationen und können sich im bestehenden institutionellen Geflecht nicht etablieren. Dies führt zu fehlender „Ownership“ des Trägers, wodurch die erreichten Erfolge nach Beendigung eines Vorhabens weitgehend verloren gehen. Aus diesem Grund spielt gerade in der Berufsbildungszusammenarbeit der menschliche Faktor eine besondere Rolle: Ohne eine kritische „Masse“ von qualifiziertem und engagiertem Personal ist weder eine überzeugende Performanz innerhalb des Vorhabens (z. B. durch kompetente Projektsteuerung und Maßnahmenumsetzung) noch eine diplomatisch geschickte Einbindung anderer Organisationseinheiten und/oder Netzwerkteilnehmenden (und damit eine dauerhafte institutionelle Verankerung) möglich. In der Regel wird die Bereitstellung entsprechender personeller Ressourcen von Dritten auch als Signal für die Bedeutung angesehen, welche die Trägerorganisation dem Vorhaben beimisst, sodass dieses in der Außendarstellung entsprechend aufgewertet wird.

Die im Modell dargestellten Erfolgsfaktoren beruhen auf einer Vielzahl von Berufsbildungsevaluationen aus den letzten 40 Jahren, in denen ein enormer konzeptioneller Wandel stattfand. Sie umfassen vielfältige unterschiedliche Projekte und Programme und wurden in den verschiedensten soziokulturellen Kontexten und politischen Systemen durchgeführt. Dies gibt Anlass zu der Vermutung, dass diese Kerngrößen eine hohe zeitliche und regionale Sta-

bilität aufweisen. Es wäre deshalb wünschenswert, wenn bei der Planung und Implementation neuer Projekte und Programme der Berufsbildungszusammenarbeit dieses Schlüsselgrößenmodell Beachtung finden würde.

Literatur

- BOLAND, Angela; CHERRY, Gemma M.; DICKSON, Rumona: Carrying out a systematic review as a master's thesis. In: BOLAND, Angela; CHERRY, Gemma M.; DICKSON, Rumona (Hrsg.): Doing a systematic review. A Student's Guide. 2. Aufl. London 2017, S. 1–21
- CASPARI, Alexandra: Lernen aus Evaluierungen. Meta-Evaluation & Evaluations-synthese von InWEnt-Abschlussequalierungen 2009. Frankfurt 2009
- CASPARI, Alexandra; BARBU, Ragnhild: Wirkungsequalierungen: Zum Stand der internationalen Diskussion und dessen Relevanz für Evaluierungen der deutschen Entwicklungszusammenarbeit. Evaluation Working Papers. Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. Bonn 2008
- COOPER, Harris: Research synthesis and meta-analysis. A step-by-step approach. 5. Aufl. Duke University 2017
- COOPER, Harris; HEDGES, Larry V.: Research synthesis as a scientific process. In: COOPER, Harris; HEDGES, Larry V.; VALENTINE, Jeffrey C. (Hrsg.): The handbook of research synthesis and meta-analysis. 2. Aufl. New York 2009, S. 3–17
- DEGEVAL – GESELLSCHAFT FÜR EVALUATION E. V. (Hrsg.): Standards für Evaluation. Erste Revision 2016. Mainz 2016 – URL: https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf (Zugriff: 31.03.2020)
- EULER, Dieter: Gemessenes und Angemessenes – Berufsbildungsforschung auf der Suche nach einem Profil... In: WEISS, Reinhold; SEVERING, Eckart (Hrsg.): Multidisziplinär – praxisorientiert – evidenzbasiert: Berufsbildungsforschung im Kontext unterschiedlicher Anforderungen. Bonn 2018, S. 30–53 – URL: <https://www.bibb.de/veroeffentlichungen/de/publication/show/9028> (Zugriff: 02.07.2020)
- FRIEDAG, Herwig R.; SCHMIDT, Walter: Balanced Scorecard. 4. Aufl. Freiburg 2011
- GLEICH, Ronald (Hrsg.): Balanced Scorecard. Best-Practice-Lösungen für die Unternehmenssteuerung. Freiburg 2012
- KAPLAN, Robert S.; NORTON, David P.: Balanced Scorecard – Strategien erfolgreich umsetzen. Stuttgart 1997
- KAPLAN, Robert S.; NORTON, David P.: Die strategiefokussierte Organisation: Führen mit der Balanced Scorecard. Stuttgart 2001
- LERNER, David: Modernization. Social Aspects. In: International Encyclopaedia of the Social Sciences (1968) 10, S. 386–402
- MANTEN, Arie A.: Scientific Literature Review, Scholarly Publishing (1973) 5, S. 75–89
- MERTENS, Donna: Research methods in education and psychology: integrating diversity with quantitative & qualitative approaches. University of Michigan 1998

- ORATA, Pedro T.: Evaluating Evaluation. In: Journal of educational research 33 (1940) 9, S. 641–661
- SCRIVEN, Michel: Evaluation Thesaurus. 4. Aufl. Newbury Park, London, New Delhi 1991
- SEVERING, Eckart; WEISS, Reinhold: Qualitätsentwicklung in der Berufsbildungsforschung: Plädoyer für eine neue Diskussion. In: SEVERING, Eckart; WEISS, Reinhold (Hrsg.): Qualitätsentwicklung in der Berufsbildungsforschung. Bonn 2012, S. 5–14 – URL: <https://www.bibb.de/veroeffentlichungen/de/publication/show/6990> (Zugriff: 02.07.2020)
- STOCKMANN, Reinhard: Die Nachhaltigkeit von Entwicklungsprojekten. Eine Methode zur Evaluierung am Beispiel von Berufsbildungsprojekten. Wiesbaden 1992
- STOCKMANN, Reinhard: Die Wirksamkeit der Entwicklungshilfe. Eine Evaluation der Nachhaltigkeit von Programmen und Projekten der Berufsbildung. Opladen 1996
- STOCKMANN, Reinhard: Handbuch zur Evaluation. Eine praktische Handlungsanleitung. Münster 2007
- STOCKMANN, REINHARD: Ziele, Wirkungen und Erfolgsfaktoren der deutschen Berufsbildungszusammenarbeit. In: Gessler, Michael; Fuchs, Martina; Pilz, Matthias (Hrsg.): Konzepte und Wirkungen des Transfers Dualer Berufsausbildung. Wiesbaden 2019, S. 121–162
- STOCKMANN, Reinhard; KOHLMANN, Uwe: Transferierbarkeit des Dualen Systems. Eine Evaluation dualer Ausbildungsprojekte in Entwicklungsländern. Berlin 1998
- STOCKMANN, Reinhard; LEICHT, René: Implementationsbedingungen eines kooperativen Ausbildungssystems in Ägypten. Berlin 1997
- STOCKMANN, Reinhard; MEYER, Wolfgang (Hrsg.): Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven. Münster 2017
- STOCKMANN, Reinhard; MEYER, Wolfgang: Chinas Berufsbildung im Wandel: 30 Jahre Entwicklungszusammenarbeit mit der Hanns-Seidel-Stiftung. Münster 2017
- STOCKMANN, Reinhard; MEYER, Wolfgang: Evaluation. Eine Einführung. 2. Aufl. Opladen, Toronto 2014
- STOCKMANN, Reinhard u. a.: Wirksamkeit deutscher Berufsbildungszusammenarbeit. Ein Vergleich staatlicher und nicht-staatlicher Programme in der Volksrepublik China. Wiesbaden 2000
- STOCKMANN, Reinhard; SILVESTRINI, Stefan (Hrsg.): Metaevaluierung Berufsbildung. Ziele, Wirkungen und Erfolgsfaktoren der deutschen Entwicklungszusammenarbeit. Münster 2013
- UMAN, Lindsay S.: Systematic reviews and meta-analyses. In: Journal of the Canadian Academy of Child and Adolescent Psychiatry, 20 (2011) 1, S. 57–59
- WIDMER, Thomas: Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern, Stuttgart, Wien 1996

Marc Beutner

► Evaluation in der Berufsbildung: das Dreischalenmodell als Grundlage

Evaluation kommt in der beruflichen Bildung eine zunehmende Bedeutung zu, die sich in generellen Aufgaben niederschlägt. Nachfolgend wird ein generelles Evaluationsverständnis bestimmt und anschließend das Dreischalenmodell der systematisch-analytischen Berufsbildungsevaluation vorgestellt. Dabei werden sein Kern mit den beinhalteten Betrachtungsebenen sowie die einzelnen Schalen eingehender betrachtet, um eine Evaluationsstruktur für Berufsbildung zu erhalten. Exemplarisch werden am Ansatz des *Design-Based-Research* die Möglichkeiten von Berufsbildungsevaluation im Spannungsfeld zwischen Gestaltung und Implementation sowie Forschung aufgezeigt. Dies wird anschließend anhand von Projektbeispielen aus der Berufsbildungsforschung für die empirische Forschung und die Entwicklungsforschung sowie deren Zusammenspiel illustriert.

1 Relevanz von Evaluation in der Berufsbildung

In jüngerer Zeit kann ein deutlicher Bedeutungsgewinn von Evaluation in der Berufsbildung festgestellt werden. Dieser Bedeutungsgewinn zeichnet sich auf verschiedenen Ebenen ab: Auf der Ebene der Durchführung von Berufsbildungsangeboten, etwa bei Kursen von Kammern und Weiterbildungsanbietern, aber auch in betrieblichen Bildungsangeboten, werden neben Kurzbefragungen zu Organisation sowie zu didaktischer Aufbereitung und zu Dozierenden verstärkt Lehrendenbefragungen und Feedbackrunden integriert, um eine Qualitätsverbesserung voranzutreiben. Auch in berufsbildenden Schulen werden inzwischen häufig Rückmeldungen von Schülerinnen und Schülern zum Unterricht sowohl in Feedbackgesprächen als auch mittels kurzer Fragebögen eingeholt. Auf der Durchführungsebene wird diese Art von Evaluation im Sinne einer Bewertung zum Teil aus Informationsgründen, zum Teil aber auch aus Legitimationsgründen vorgenommen, z. B. im Rahmen der Schulaufsicht und Qualitätssicherung. Diese Legitimationsgründe gepaart mit Nutzengründen finden sich auch bei der Einbettung von Evaluationen in der Planungs- und Organisationsebene, etwa bei schulischen Evaluationen zu Bildungsgängen oder zum Verbleib der Schülerinnen und Schüler. Im Rahmen von betrieblicher Aus- und Weiterbildung kann auch eine organisatorische Evaluation durchgeführt werden, die sich etwa auf den Prozess der Kursplanung oder auf

den Nutzen für ein Unternehmen beziehen kann – auch mit Blick auf ganze Institutionen und auf das Berufsbildungssystem. Neben der Tatsache, dass in immer mehr Bereichen der Berufsbildung eine Bewertung und eingehendere Betrachtung vorgenommen wird, kann auch in vielen dieser Bereiche eine deutlich elaboriertere Vorgehensweise ausgemacht werden. In den letzten Jahren werden beispielsweise an immer mehr Berufskollegs Evaluationsteams gebildet, wie etwa am Georg-Simon-Ohm Berufskolleg in Köln, das ein eigenes Evaluationsportal eingerichtet hat (vgl. GEORG-SIMON-OHM BERUFSSKOLLEG 2019). Dort werden z. B. einfache Methoden und Hilfestellungen für Selbstevaluationen im Unterricht, Lehrer- und Schülerfeedbackbögen, ein computergestütztes Auswertungstool, eine Beispielauswertung und eine Evaluationszielscheibe zur Lernsituationsreflexion zur Verfügung gestellt. Einen übergeordneten Blickwinkel einnehmend kann „für die Berufsbildung festgestellt werden, dass Evaluationen zunehmen und zum einen systematischer durchgeführt werden und zum anderen mehr Multiperspektivität aufweisen, als dies noch vor einigen Jahren der Fall war“ (BEUTNER 2019, S. 121).

Zusammenfassend kann festgehalten werden, dass ein Bedeutungsgewinn von Evaluationen in der Berufsbildung auszumachen ist und Berufsbildungsevaluation schon seit längerer Zeit immer wieder thematisiert wird (vgl. BEUTNER 2009), aber nicht immer klare Orientierungen an Evaluationsmodellen zum Einsatz kommen (vgl. BEUTNER 2007; 2018; 2019). Sofern eine solche Orientierung erfolgt, bedeutet das noch nicht, dass diese stets systematisch umgesetzt wird.

2 Verständnis von Evaluation und ihren Aufgaben

Michael Scriven definierte bereits im Jahre 1991 Evaluation in folgender Form:

„Evaluation refers to the process of determining the merit, worth, or value of something, or the product of the process [...]. The evaluation process normally involves some identification of relevant standards of merit, worth, or value; some investigation of the performance of the evaluands on these standards; and some integration or synthesis of the results to achieve an overall evaluation or set of associated evaluations.“ (SCRIVEN 1991, S. 139)

Er spezifiziert dies 2007 mit Blick auf die zur sinnvollen Evaluation notwendige Expertise, indem er ausführt:

„Evaluation is the process of determining merit, worth, or significance; an evaluation is a product of that process. Professional evaluation is evaluation done in a systematic and objective way with a degree of expertise that requires extensive specific training or learning.“ (SCRIVEN 2007)

Evaluation ist durch ein Zusammenspiel von (a) Personen, die die Evaluation durchführen, den Evaluationssubjekten, den (b) Evaluationsobjekten, also den Programmen, Prozessen, Produkten, Personen(-gruppen) und Maßnahmen etc., die in der Evaluation untersucht werden, und den (c) Bewertungsprozessen und Bewertungen gekennzeichnet (vgl. BEUTNER 2018, S. 37).

Achtenhagen verwies bereits 1995 darauf, dass komplexe berufsbildende Lehr-/Lernarrangements ein Evaluationsobjekt in der Berufsbildung darstellen können (vgl. ACHTENHAGEN 1995). Zudem verdeutlichte Buschfeld bereits 2002 in seinen Überlegungen zu Konditionen beruflicher Bildungsgänge, dass neben der Kondition des Curriculums und der Kondition der Ressourcen die dritte wesentliche Kondition die Evaluation sei (vgl. BUSCHFELD 2002). Auch Kranert/Kremer nehmen die Bedeutung von Evaluation in ihren Überlegungen zum Qualitätsmanagement in Innovationsarenen auf (vgl. etwa KRANERT/KREMER 2017, S. 15f., 23f.).

Dabei sind jeweils unterschiedliche Perspektiven auf den Evaluationsbegriff möglich, etwa aus der Alltagsperspektive, der Modellperspektive, der Prozessperspektive, der Ergebnisperspektive, der Wertperspektive und der institutionellen Perspektive (vgl. BEUTNER 2018, S. 47). Mit Blick auf die Berufsbildungsevaluation sind dabei verschiedene generelle Aufgaben dieser Evaluationen auszumachen. So werden Berufsbildungsanalysen und Berufsbildungsdokumentationen vorgenommen, Berufsbildungsdesigns umgesetzt und untersucht, Berufsbildungsaufgaben erfüllt und hinsichtlich des Erfüllungsgrades betrachtet sowie Berufsbildungsbewertungen vorgenommen. In diesem Kontext finden sich konkrete didaktische, pädagogische, curriculare, berufsbezogene, arbeitsmarktbezogene, organisatorische, technische, regulatorische und institutionelle Aufgaben, die jeweils zu berücksichtigen sind. Kosten-Nutzen-Relationen stehen dabei genauso im Blick wie Qualitäts-, Effizienz- und Effektivitätsfragen.

Um grundlegende Eigenschaften von Evaluationen zu fixieren, wurden von der DeGEval, der Deutschen Gesellschaft für Evaluation, Standards für Evaluationen festgeschrieben (vgl. etwa DEGEVAL 2016). Als übergeordnete Standardbereiche zählen hierbei (a) Nützlichkeit, (b) Durchführbarkeit, (c) Fairness und (d) Genauigkeit, die jeweils durch die einzelnen Standards spezifiziert werden (vgl. BEUTNER 2019; DEGEVAL 2016).

Als Beispiele für Evaluationsbereiche in der Berufsbildung lassen sich etwa die Analyse von Lernangeboten und Lehrkonzeptionen, die Analyse von Berufsbildungsinstitutionen, die Analyse von berufsbildungspolitischen Maßnahmen und Programmakzeptanz, die Analyse von (Vorbereitungs-, Aus- und Weiterbildungs-)Maßnahmen sowie die Analyse von Transfermöglichkeiten nennen, die in der aktuellen Umsetzung von Berufsbildungsevaluationen sowie deren Diskussion alle eine Rolle spielen.

Systematische Berufsbildungsevaluation bezieht sich auf den Prozess der Analyse und Bewertung eines Evaluationsobjekts bzw. der Effekte und Wirkungen von Maßnahmen oder Programmen aus Berufsbildungskontexten in Wirtschaft, Schule und Gesellschaft. Sie erfolgt als Evaluierungsprozess formativ, summativ oder askriptiv und beinhaltet typischer-

weise die Identifizierung relevanter Standards und Kriterien sowie ein nachvollziehbares systematisch-methodisches Vorgehen in der Regel unter Nutzung und Einbindung unterschiedlicher Perspektiven: „Aktuelle Evaluationen in der Berufsbildung werden oftmals mit Blick auf mehrere Stakeholder konzipiert und umgesetzt“ (BEUTNER 2019, S. 121).

3 Evaluationsgrundlage: Dreischalenmodell der systematisch-analytischen Berufsbildungsevaluation

Das Dreischalenmodell der systematisch-analytischen Berufsbildungsevaluation (vgl. BEUTNER 2018), das im Folgenden in komprimierter Form vorgestellt wird, ist ein Strukturmodell, das eine Orientierung und Hilfestellung bieten soll, wenn Evaluationen in der Berufsbildung zu konzipieren und umzusetzen sind. Es geht dabei darum, eine robuste Struktur und eine Heuristik an die Hand zu bekommen, die zum einen theoretisch basiert ist, zum anderen aber Bereiche aufzeigt, in denen Entscheidungen und Verortungen im Rahmen der Evaluation vorgenommen werden sollten. Das Dreischalenmodell arrangiert drei Schalen um einen Kern, in dem eine Betrachtung von Evaluationsobjekten der Berufsbildung auf mehreren Ebenen angeregt wird und dem Gedanken der Multiperspektivität Rechnung getragen werden soll.

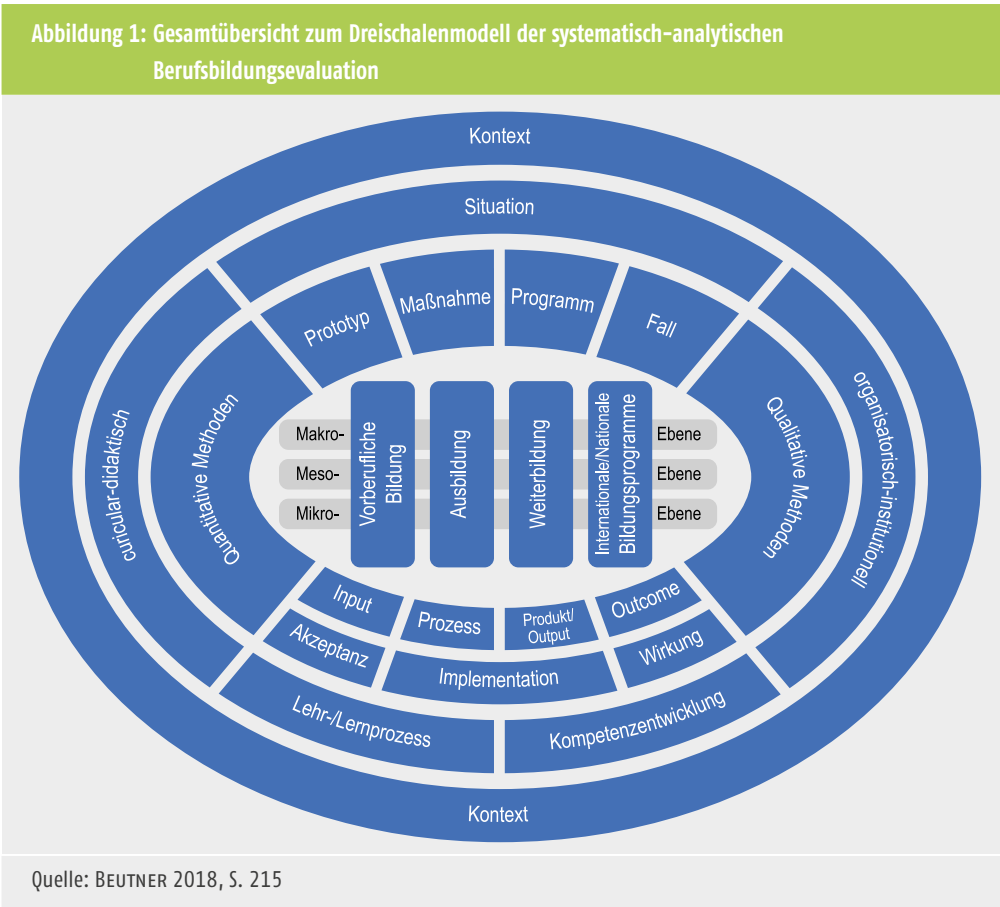
Generell orientieren sich Evaluationen im Rahmen des Dreischalenmodells an Kromreys drei Zielen von Evaluation (vgl. KROMREY 2006, S. 102ff.). Nach Kromrey kann ein erstes Ziel die Bildung eines Werturteils sein. Es fokussiert Wirkungen im Rahmen von Wirkungsanalysen, welche die Effekte von Maßnahmen o. Ä. untersuchen. Zweitens können Umsetzungen bzw. Implementationen im Rahmen der Implementationsforschung, die eine systematische Betrachtung von Planung, Durchführung und Umsetzung beinhaltet, untersucht werden. Als dritte mögliche Zielsetzung betrachtet Kromrey die Akzeptanz von Maßnahmen. Dies geschieht im Rahmen der Akzeptanzforschung, die die Zustimmung oder Ablehnung von Maßnahmen oder anderen Evaluationsobjekten durch Personen oder Personengruppen untersucht.

Evaluationen im Rahmen des Dreischalenmodells sind oft im Sinne Scrivens (vgl. SCRIVEN 1991) summativ oder formativ oder in Kombination dieser Formen angelegt. Aber auch evolutive und askriptive Vorgehensweisen sind möglich. Summative Evaluationen betrachten dabei ein Endergebnis, nachdem eine Intervention, Maßnahme o. Ä. bereits abgeschlossen ist, während formative Evaluationen begleitend zur Maßnahme durchgeführt werden und mit ihren Ergebnissen und deren Aufnahme durch die Personen, die eine zu evaluierenden Maßnahme umsetzen, durchaus die künftige Ausgestaltung beeinflussen.

Bei askriptiven Evaluationen (vgl. SCRIVEN 2004a, S. 167ff.; SCRIVEN 2004b, S. 17) stehen der Wert und der Nutzen eines Evaluationsobjektes im Blickpunkt, wobei bilanzierende Schlussfolgerungen getroffen werden und eine unmittelbare Entscheidungsfindung unterstützt und angestrebt wird, es geht jedoch nicht um Verbesserungsgedanken. Die Nutzenbestimmung und Nutzenzuschreibung wird dabei Evaluationszweck. Scriven macht dies da-

ran deutlich, dass etwa ein Historiker Napoleons Nutzung der Kavallerie bewerten könnte, er aber in keiner Weise Überarbeitungen, Verbesserungen oder Entwicklungen intendieren würde (vgl. SCRIVEN 2004b, S. 17).

Im Rahmen von evolutiven Evaluationen im Sinne Michael Quinn Pattons (vgl. PATTON 2010) steht die Gestaltung im Vordergrund und der Evaluationsgegenstand bzw. das Evaluationsobjekt wird mit Unterstützung der Evaluation entwickelt. Dieses Vorgehen ist etwa im Rahmen von *Design-Based-Research* (DBR)-Ansätzen bisweilen beobachtbar und sinnvoll umzusetzen.



3.1 Der Kernbereich des Modells

In seinem Kern stellt das Dreischalenmodell die ebenenbezogene und ebenenübergreifende Betrachtung von Maßnahmen und Programmen im Bereich der Vorberuflichen Bildung (VB), der Ausbildung (Ausb.) und der Weiterbildung (WB) ins Zentrum (vgl. BEUTNER 2018, S. 219).

In der Berufsbildung kann dabei jeweils ein unterschiedlicher Fokus auf die zu betrachtenden Ebenen gelegt werden. Dies ermöglicht es dem Nutzer bzw. der Nutzerin, auf verschiedene Weisen eine Mehrperspektivität einzubringen, die nicht allein eine zielgruppenspezifische Mehrperspektivität darstellt. Untersuchungen können dabei im allereinfachsten Falle auf lediglich einer Ebene in einem Bereich (VB, Ausb. oder WB) durchgeführt werden. Es können auch Analysen auf einer Ebene in mehreren Bereichen (z. B. auf der Makroebene in Ausb. und WB) oder auf mehreren Ebenen in einem Bereich (z. B. auf der Mikro- und Mesoebene im Bereich Ausb.) erfolgen. Jedoch kann auch auf mehreren Ebenen in mehreren Bereichen (z. B. auf der Mikro- und Mesoebene in den Bereichen VB und Ausb.) eine Untersuchung umgesetzt werden.

In einer Deutung nach Sloane kann eine bildungsmaßnahmenbezogene Sicht auf die Ebenen herangezogen werden. Dabei wird die Mikroebene mit Blick auf den Unterricht und dortige Lernsettings und Lernsituationen gedeutet, die Mesoebene mit Blick auf die Organisation, sprich die Arbeit und Kollaboration im Bildungsgang und im Bildungsgangteam, sowie die Makroebene mit Blick auf die curricularen Vorgaben, wie etwa den Lehrplan (vgl. SLOANE 2003, S. 7).

Alternativ kann im Modell eine Deutung der Ebenen im Sinne der Soziologie erfolgen, wobei die Mikroebenenbetrachtung dann den Fokus auf eine Individual- oder Kleingruppensicht legt, die Mesoebenenbetrachtung sich der Organisation als ganzer und die Makroebenenbetrachtung sich dem gesellschaftlichen Fokus widmet (vgl. BEUTNER 2018, S. 219f.).

Zudem haben Nutzer/-innen des Modells auch die Möglichkeit, die Untersuchungsebenen in Bezug auf Programme und Maßnahmen zu deuten. In diesem Falle erfolgt auf der Mikroebene die Maßnahmendurchführung, auf der Mesoebene die Betrachtung des Programm- bzw. Maßnahmenrahmens und auf der Makroebene die Untersuchung der Programm- bzw. Maßnahmenvorgaben bzw. der Zielsetzungen (vgl. BEUTNER 2018, S. 219f.).

Nicht zuletzt kann der Nutzer oder die Nutzerin sich auch dazu entscheiden, die Ebenen anhand des einbezogenen Reflexionscharakters zu deuten. Hier würde auf der Mikroebene eine objektenebenbezogene Untersuchung erfolgen, auf der Mesoebene eine theoriebezogene Reflexion der Informationen zur Objektebene umgesetzt sein und auf der Makroebene eine Reflexion der Informationen zur Theorie vorgenommen werden, sprich eine Metatheorieebene (vgl. BEUTNER 2018, S. 219f.).

Nachdem der Nutzer oder die Nutzerin des Modells damit den Kern der Untersuchung über die Festlegung der zu betrachtenden Bereiche und Ebenen bestimmt hat, werden auf den Schalen weitere Konkretisierungen und Orientierungen gegeben und damit Felder aufgezeigt, in denen Entscheidungen zu treffen sind.

3.2 Die erste, innere Schale des Modells

Auf der ersten Schale, der inneren, wird eine methodische Konkretisierung vorgenommen, um eine systematische Herangehensweise an die Evaluation zu gewährleisten. Dabei können qualitative und/oder quantitative Methoden zum Einsatz kommen. Im Falle einer Verbindung der beiden Formen ist intensiv über die Methodentriangulation (vgl. etwa MAYRING 2001 oder STECKLER u. a. 1992) etwa im Sinne eines Vorstudien-, Verallgemeinerungs-, Vertiefungs- oder eines tatsächlichen Triangulationsmodells zu reflektieren.

Die Entscheidung, welche Methoden im qualitativen und/oder quantitativen Bereich Anwendung finden können, hängt nicht zuletzt davon ab, ob die Evaluation einer Maßnahme fokussiert werden soll, ein Prototyp gestaltet und evaluiert werden soll, eine Programmevaluation vorzunehmen ist oder aber eine Fallstudie anhand eines oder mehrerer Fälle mit spezifischen Rahmenbedingungen zu untersuchen ist. Zudem spielt bei der Methodenüberlegung eine wesentliche Rolle, welchen Blick Evaluatorinnen und Evaluatoren mit der zu wählenden Methodik auf das oder die Evaluationsobjekte werfen möchten. Hier kann

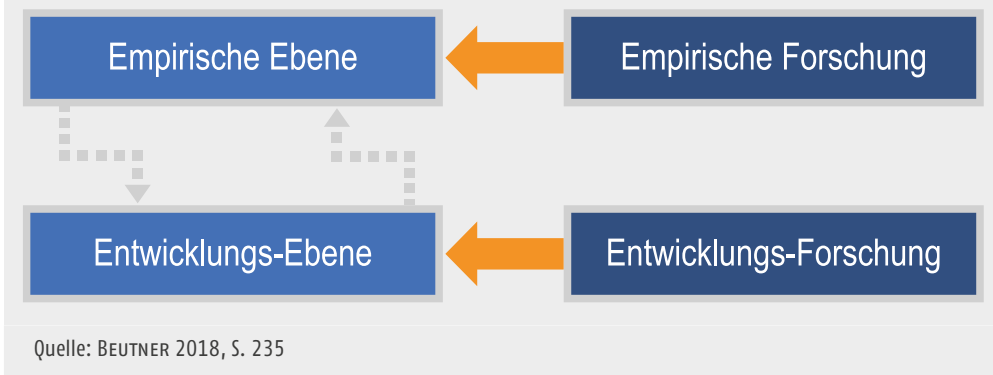
- a) eine Inputanalyse und die oft damit verbundene Betrachtung von Akzeptanz im Sinne der Akzeptanzforschung Kromreys erfolgen,
- b) ein Fokus auf die Kromrey'sche Implementationsforschung mit Prozessanalysen und/oder einer Produkt- bzw. ergebnisbezogenen Perspektive im Sinne einer Outputanalyse gerichtet werden oder
- c) eine Wirkungsforschung nach Kromrey mit Blick auf den Outcome und damit eine ggf. zeitlich längere Perspektive, nicht zuletzt mit Blick auf Transfer und Nachhaltigkeit, fokussiert werden.

Je nach Entscheidungen in dieser Schale, sind die methodischen Instrumente und Verfahrensweisen an diesen auszurichten. Ziel ist es, adäquate Methoden zur Zielerreichung im Rahmen der Evaluation auszumachen, eingehender zu definieren und entsprechend zum Einsatz kommen zu lassen.

Als ein methodisches Beispiel sei hier der multimethodische Ansatz mit Kombination qualitativer und quantitativer Umsetzungen im Rahmen des DBR dargestellt. Dies dient der Illustration und meint damit nicht, dass dies der einzig gangbare oder aber etwa der richtige Weg im Rahmen der methodischen Gestaltung sei.

Sofern etwa in Schale 1 ausgemacht wird, dass beispielsweise eine Prototypenentwicklung und die Evaluation dieses Prototyps sowie des dazugehörigen Entwicklungsprozesses im Blickpunkt der Evaluation stehen, liegt der Fokus auf einer Implementationsforschung nach Kromrey. Im DBR-Ansatz wird unterschieden zwischen einer empirischen Ebene, auf der die empirische Forschung ansetzt, und einer Entwicklungsebene, auf der die Entwicklungsforschung ansetzt.

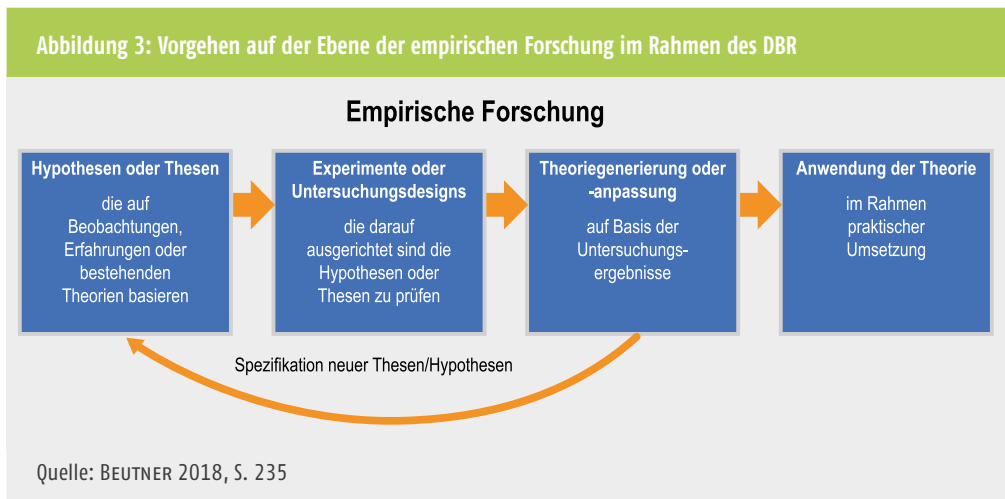
Abbildung 2: Forschung auf empirischer Ebene und Entwicklungsebene im DBR-Ansatz



Wesentliches Charakteristikum des DBR-Ansatzes ist hierbei, dass die gewonnenen Erkenntnisse auf der empirischen Ebene die Entwicklungsebene beeinflussen und die Erfahrungen auf der Entwicklungsebene ebenfalls Rückwirkungen auf die weiteren empirischen Untersuchungen haben.

Auf der Ebene der empirischen Forschung werden im Rahmen der Evaluation Thesen und ggf. Hypothesen (je nach qualitativer oder quantitativer Forschungsausrichtung) mit Blick auf das Erkenntnisinteresse im Kontext des Prototyps aufgestellt, welcher im Rahmen der Entwicklungsforschung heranreift und spezifiziert wird. Die Thesen und Hypothesen basieren auf Beobachtungen, auf eigenen Erfahrungen oder Erfahrungen anderer oder aber auf bestehenden Theorien und Modellen. Anschließend werden Experimente oder Untersuchungsdesigns aufgestellt, die darauf ausgerichtet sind, die vorher bestimmten Thesen oder Hypothesen zu prüfen. Dabei geht es nicht zwingend um eine Verifikation der Hypothesen, sondern im Popper'schen Sinne (vgl. POPPER 2005) um die Möglichkeit der Falsifikation (vgl. POPPER 1989, S. 82–85).

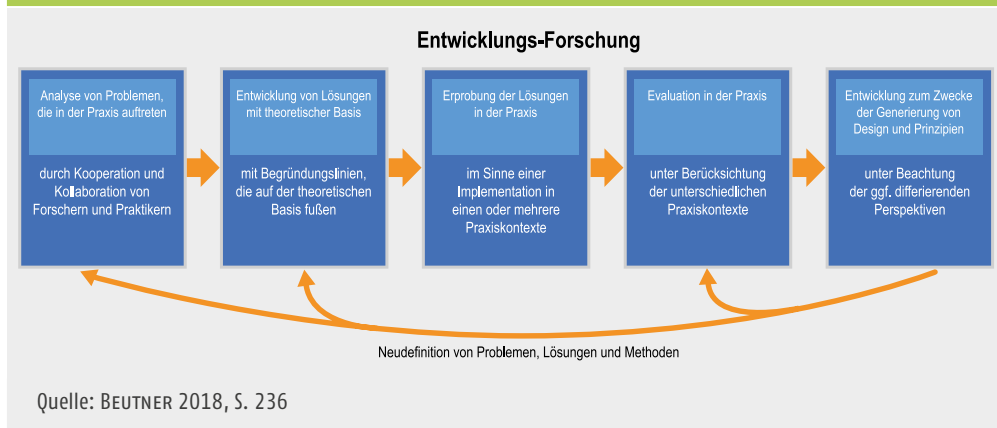
Aufbauend auf den gewonnenen Erkenntnissen aus dem empirischen Forschungsprozess, erfolgt eine Theoriegenerierung oder eine Anpassung bestehender Theorien. Auch können bisherige Theorien und Modelle sich als falsch oder fehlerhaft herausstellen, was oftmals zur Generierung neuer Vorstellungen und Theorien führt. Zudem können die Ergebnisse auch zu Spezifizierungen bestehender Thesen und Hypothesen oder aber zur Generierung neuer Thesen und Hypothesen führen, die dann wiederum zu überprüfen sind. Ziel ist es, eine Anwendung der Theorie im Rahmen praktischer Umsetzung zu unterstützen.



Auf der Entwicklungsebene erfolgt die eigentliche Ausarbeitung des Prototyps. Hier ist somit der Design-Prozess zu finden, auf dem die zuvor dargelegte empirische Forschung basiert. Analysiert werden dazu zunächst die Herausforderungen und Probleme, die in der Praxis zu finden sind und für die der zu entwickelnde Prototyp eine Lösung darstellen soll. Die Analyse erfolgt in einer Kooperation und Kollaboration von Forschenden und Praktikern bzw. Praktikerinnen, damit sich beide Seiten der Ausgangsproblematik bewusst sind und erkennen, wie diese jeweils beschrieben und durch welche Elemente sie gekennzeichnet wird.

Anschließend erfolgen erste Überlegungen zur Entwicklung einer Lösung, wobei hier stets eine theoretische Basis zum Ausgangspunkt genommen wird. Das bedeutet, dass die Begründungen für die Wahl des Lösungsweges einer Reflexion bedürfen und in der Regel eben nicht auf einem reinen *Trial-and-Error*-Prinzip beruhen. Diese Lösungsansätze werden in einem oder mehreren Praxiskontexten in Teilen oder als Ganzes implementiert, um die Güte der Lösung bzw. den Lösungsfortschritt erkennen zu können. Dies geschieht über eine Evaluation der Praxis, wobei die verschiedenen Praxiskontexte mit ihren unterschiedlichen Zielgruppen, Rahmenbedingungen und Ansprüchen mehrperspektivisch aufgenommen werden. Ziel ist hierbei neben der eigentlichen Prototypenentwicklung als stabile erste Lösung auch die Entwicklung und Reflexion der Entwicklungsergebnisse zum Zwecke der Generierung von Design-Prinzipien. Diese entstehen unter Beachtung der verschiedenen Perspektiven und deren Unterschiedlichkeit, was Weiterentwicklungen und Transferüberlegungen unterstützt. Anhand des Standes der Entwicklung der Design-Prinzipien kann es zu Rückwirkungen auf die Definition und Analyse des Ausgangsproblems, auf den Prototypen als aktuelle Lösung sowie auf dessen Implementationsevaluation kommen. Somit wird über die Neudefinition von Problemen, Lösungen und Methoden ein zirkulärer Entwicklungsprozess angestoßen, an dessen Ende ein stabiler Prototyp sowie stabile Design-Prinzipien stehen.

Abbildung 4: Vorgehen auf der Ebene der Entwicklungsforschung im Rahmen des DBR



3.3 Die zweite, mittlere Schale des Modells

Die zweite Schale ist durch berufsbildungsspezifische Besonderheiten gekennzeichnet, die es als Mindeststandard zu beachten gilt. Dabei ist die Zuordnung nicht dauerhaft vollständig, sondern stets vor dem Hintergrund der aktuellen Lage und Gegebenheiten in der Berufsbildung zu reflektieren und zu ergänzen.

Wesentliche Kennzeichen der Berufsbildung, von denen das Modell ausgeht, sind die besonderen curricular-didaktischen Rahmenbedingungen sowie die organisatorisch-institutionellen Rahmenbedingungen in Wirtschaftskontexten. Dies konkretisiert sich nicht zuletzt über die berufliche Kompetenzentwicklung und das Ziel beruflicher Handlungskompetenz sowie über die zugehörigen beruflichen Lehr-/Lernprozesse. So sind etwa die Bedeutung der Situationsorientierung mit der Anknüpfung an betriebliche und/oder wirtschaftliche Situationen sowie die curriculare Abbildung solcher Situationen ein Ausdruck der zweiten Schale. Beispiele hierfür sind die Abbildung von betrieblichen Handlungssituationen, Lernfeldern und Lernsituationen in der lernfeldorientierten Didaktik und lernfeldorientierten Curricula. Auch die Abbildung authentischer betrieblicher Situationen im Rahmen des *E-Learning* und der *Serious-Game*-Gestaltung für berufliche Kontexte gehören dazu. Hier wird im Rahmen der Evaluation ermittelt, inwieweit sie diesen spezifischen Besonderheiten der Berufsbildung gerecht wird und wie sie sich darin platziert.

3.4 Die dritte, äußere Schale des Modells

Mit der äußeren Schale werden die Einflussfaktoren und Kontextvariablen aufgenommen, die ggf. Auswirkungen haben, jedoch nicht direkt den unmittelbaren beruflichen Besonderheiten entspringen wie etwa die gesellschaftlich-politische Lage, kulturelle Einbettung oder europäische Vorgaben. Hier findet eine Kontextevaluation statt.

Mit Blick auf die gesellschaftlich-politische Lage sind das jeweilige Partizipationsverständnis, die aktuellen Schwerpunktthemen und Diskussionen in der Gesellschaft, die Bedeutungseinschätzung von Bildungspolitik und beruflicher Bildung sowie die finanzielle Ausstattung der Bildung von Bedeutung. Zudem stehen auch Fragen wie der Umgang mit den Geschlechtern und deren Gleichstellung, der Umgang mit Diskriminierungen, Fragen der Gewaltprävention, die Berücksichtigung von Benachteiligungen in der Gesellschaft im Fokus. Weitere sozialpolitische Aspekte sind zudem die Berücksichtigung sozial schwacher Gruppen, von Sozialleistungen und der notwendigen Infrastruktur.

Die kulturelle Einbettung hingegen betrachtet das Verhältnis von nationaler Gesellschaft sowie ihren Werten zu anderen Gesellschaften oder Zusammenschlüssen wie etwa anderen Ländern oder der Europäischen Union. Hier stehen Werte und Normen sowie Verhaltensweisen und Sichtweisen im Zentrum der Betrachtung. Mit Blick auf Internationalisierung und Globalisierung kommt diesen Aspekten eine zunehmende Bedeutung zu. So ist etwa das Verhältnis von nationaler Staatsbürgerschaft und Europäischem Staatsbürgerschaftsverständnis zu betrachten. Gesellschaftlich oder geschichtlich erworbene Aspekte wie Sprache, Religion, Nationalverständnis, Werte, Sitten und Gebräuche spielen wichtige Rollen.

Nationale und europäische sowie internationale Vorgaben und Abkommen müssen auf Gemeinsamkeiten und Unterschiede untersucht werden, da sie oftmals einen entscheidenden Rahmen für die jeweiligen Aktivitäten, Maßnahmen und Handlungen bilden.

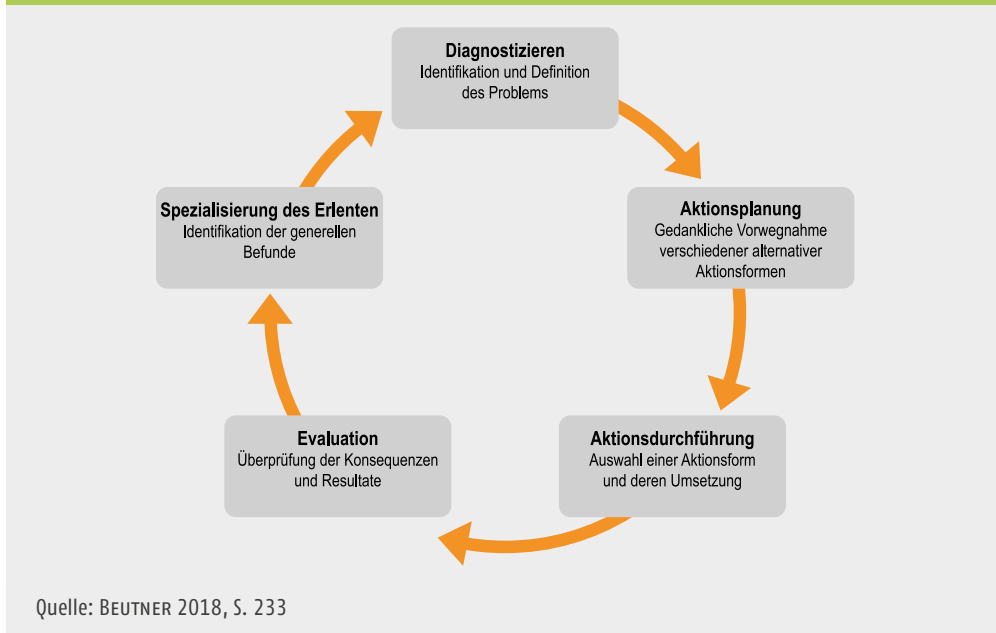
4 Einblicke in die Anwendung des Dreischalenmodells der systematisch-analytischen Berufsbildungsevaluation

Um einen Einblick in das Vorgehen mittels einer systematisch-analytischen Berufsbildungsevaluation zu geben, sollen exemplarische Einblicke in Projekte gegeben werden, die dem Gedanken des DBR folgten. Dabei soll zunächst die gestalterische Entwicklungsebene und anschließend die empirische Forschungsebene aufgezeigt werden, da mit der Darstellung der Entwicklung auch ein Einblick in die Grundidee des Projekts vermittelt werden kann.

4.1 Eindrücke aus Bereichen der Entwicklungsforschung im Rahmen von DBR-Ansätzen

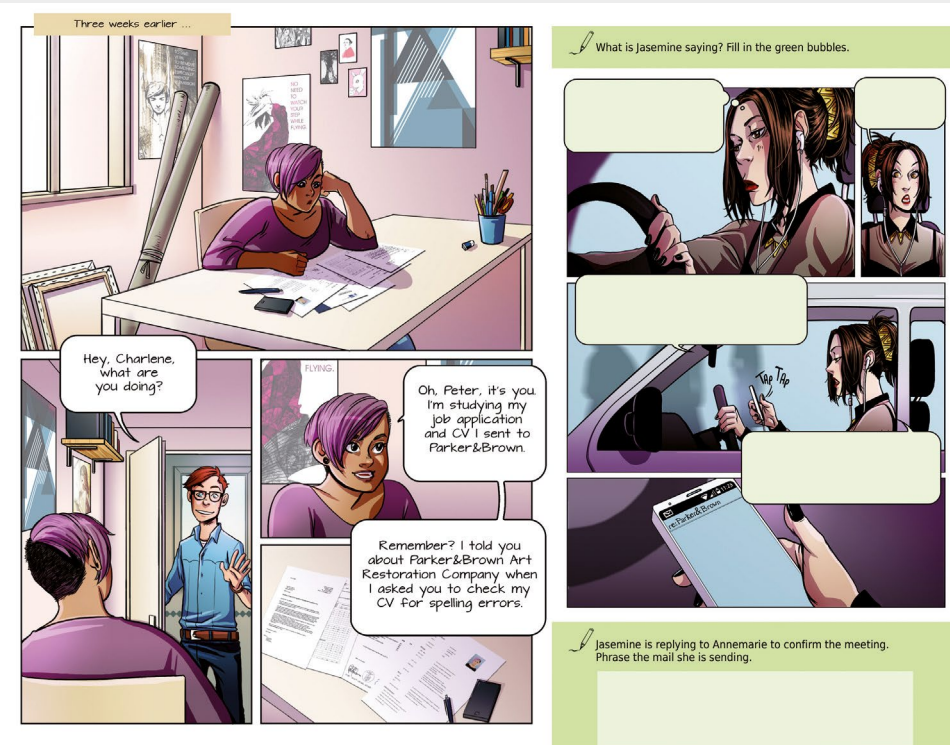
In der Entwicklungsforschung steht oftmals die Generierung eines Prototyps im Blickpunkt des Interesses.

Abbildung 5: Prototypenentwicklung unter dem Eindruck der Aktionsforschung



Mit Blick auf die Ebene der Entwicklungsforschung soll zunächst einmal das Projekt „AG-novel – Advanced Interactive Graphic Novels on Mobile Touchscreen Devices“ (vgl. BEUTNER/PECHUEL 2015) betrachtet werden, das im Rahmen des *Lifelong-Learning*-Programms mit Mitteln der Europäischen Kommission umgesetzt wurde. Ziel des Projektes „AGnovel“ war es, eine moderne angemessene Lernumgebung für das Sprachenlernen zu schaffen. Dies wird durch die Gestaltung eines Konzeptes für elektronische *Graphic Novels* für die allgemeine und berufliche Bildung realisiert. Die Projektpartner entwickelten dazu eine *Graphic Novel*, die speziell an Realsituationen des Alltags und des Berufs anknüpft und auf mobilen Endgeräten, sprich als Tablet-App umgesetzt, erprobt und evaluiert wurde. Hierbei können die Texte in den Sprechblasen der *Graphic Novel* in verschiedenen Sprachen dargestellt werden und auch bei Anklicken der Sprechblasen mittels Audio von Muttersprachlern oder Muttersprachlerinnen gesprochen werden. So kann für die Lernenden beispielsweise der Text in Englisch erscheinen, beim Tippen auf die Sprechblase jedoch zur Unterstützung ein deutsches Audio abgespielt werden. Dies ist auch in anderen Sprachkombinationen möglich. Zudem sind in die App didaktische Aufgaben eingebaut, die bei Bedarf von den Dozenten oder Dozentinnen zugeschaltet werden können. Sie sind so ausgerichtet, dass fächerübergreifender Unterricht zwischen den Fremdsprachen, der Betriebswirtschaftslehre mit Bereichen wie Kostenrechnung und Rechnungswesen, dem Fach Deutsch und Kommunikation, Kunst sowie Ethik/Religion möglich ist.

Abbildung 6: Einblicke in die entwickelte App und die Aufgaben von „AGnovel“



Three weeks earlier ...

Hey, Charlene, what are you doing?

Oh, Peter, it's you. I'm studying my job application and CV I sent to Parker & Brown.

Remember? I told you about Parker & Brown Art Restoration Company when I asked you to check my CV for spelling errors.

What is Jasmine saying? Fill in the green bubbles.

Jasmine is replying to Annemarie to confirm the meeting. Phrase the mail she is sending.

Quelle: BEUTNER/PECHUEL 2015

Im Rahmen der Entwicklungsevaluation wurden z. B. der Umgang mit der App und die didaktischen Einbettungsmöglichkeiten erhoben.

Ein zweites Beispiel sei mit dem europäisch ausgerichteten Projekt „MATH – Mountains of absolutely terrifying height“ gegeben, das mit Mitteln der Europäischen Kommission aus dem ERASMUS+-Programm umgesetzt werden konnte. Um mathematische Kompetenzen und Kenntnisse rund um den Bereich der Steigung zu vermitteln, wurde ein kompetitives Spiel entwickelt. In diesem Spiel müssen Lernende im Wettkampf gegen andere Spieler/-innen schnellstmöglich Berge erklimmen, wobei eben diese mathematischen Kompetenzen unmittelbar zum Einsatz kommen. Ein Ziel des Projektes „MATH“ ist neben der Umsetzung des Tools vorrangig eine Förderung der Motivation für Lernprozesse. Der spielbasierte Lernansatz (*game-based learning approach*) bietet Trophäen für Errungenschaften (wie Medaillen für Berge, die die Lernenden im Spiel besteigen). „MATH“ ist ein *Serious Game*, das eine positive Begegnung mit Mathematik in allgemeinbildenden schulischen und berufsschulischen Kontexten ermöglicht. Im *Challenge*-Modus wird ein asynchroner Ansatz verfolgt, der es den

herausfordernden Personen ermöglicht, zuerst den Berg zu besteigen. Lernende, die herausgefordert werden, versuchen dann, den Berg schneller zu erklimmen. Dabei können Lernende beide Kämpfer/-innen nebeneinander klettern sehen. Die Umsetzung kann komplett nacheinander oder aber in Runden erfolgen. Inhaltlich werden die Steigungen über Brüche, Steigungsdreiecke, Umrechnungen anhand des Satzes des Pythagoras usw. abgebildet.

Abbildung 7: Einblicke in die entwickelte App und die Aufgaben von „MATH“



Quelle: BEUTNER/PECHUEL 2019

Im Rahmen der Entwicklungsevaluation lag hier der Fokus auf der Einbettung der App in den schulischen und berufsschulischen Unterricht.

Ein drittes Beispiel ist die Entwicklung im innerdeutschen Projekt „NetEnquiry“, das über das Bundesministerium für Bildung und Forschung gefördert wurde. In „NetEnquiry“ wurden neuartige mobile Lernangebote in Kombination mit dem dazugehörigen Softwaretool für den Bankenbereich entwickelt, erprobt und evaluiert. Abgebildet wurden reale Handlungsabläufe und Arbeitsplatztools (etwa ein Kreditrechner, ein komplettes E-Mail-System, ein Chat-System, eine Simulation von interaktiven Gesprächen mit Kundinnen und Kunden usw.). Auch hier wurde ein kompetitiver Ansatz bei den Lernenden verfolgt, wobei die Lernenden verschiedene Rollen übernehmen (Kreditberater/-in, Entscheider/-in, Backoffice-Mitarbeiter/-in usw.), kooperativ zusammenarbeiten und gegen andere Lernendengruppen antreten. Das Projekt ermöglicht somit die Nutzung der Potenziale des mobilen Lernens im Rahmen der betrieblichen Aus- und Weiterbildung. Im Projekt „NetEnquiry“ werden realitätsnahes Lernen mit dem Einsatz mobiler Endgeräte verknüpft, Anforderungen des Arbeits-

umfeldes berücksichtigt, und es wird dazu beigetragen, komplexe betriebliche Abläufe und Systeme in ihrer Gesamtheit zu erlernen und zu verstehen.

Abbildung 8: Einblicke in das NetEnquiry-Tool



Quelle: BEUTNER/PECHUEL 2013; BEUTNER 2016

Im Rahmen der Entwicklungsevaluation wurden hier betriebliche Ablaufstudien und Tätigkeitsbeobachtungen umgesetzt, die den Authentizitätsgrad der Abbildung erhöhen, und zudem qualitative und quantitative Befragungen zu Verbesserungsmöglichkeiten im Rahmen der Entwicklung durchgeführt.

4.2 Eindrücke aus Bereichen der empirischen Forschung im Rahmen von DBR-Ansätzen

Mit Blick auf die Ebene der empirischen Forschung wurden im Projekt „AGnovel“ mehrperspektivische Untersuchungen des „AGnovel“-Prototypen sowie der „AGnovel“-*Blended Learning*-Bildungsmaßnahme umgesetzt (vgl. BEUTNER/PECHUEL 2015). Dies geschieht mit Blick auf die Tablet-App und ihre Kurseinbettung auf der Mikroebene im Bereich der beruflichen Erstausbildung sowie auf der Mesoebene mit Blick auf die organisatorische Einbettung in der Ausbildung und bei internationalen Bildungsprogrammen. Die Mikroebene wird dabei in Individualsicht bzw. Kleingruppensicht untersucht, während die Mesoebene Individualsichten zu einer Organisationssicht aggregiert. Methodisch erfolgt die Untersuchung im Projekt „AGnovel“ mittels quantitativer Befragungen bei Lernenden (n=146) und Lehrenden (n=70). Zudem wurden eine qualitative Analyse und eine Interviewstudie

durchgeführt, die die didaktischen Anforderungen und Umsetzungsmöglichkeiten des Sprachenlernens mittels interaktiver Apps untersuchten. Hierbei wurde das CIPP-Modell nach Stufflebeam (vgl. STUFFLEBEAM 2003, S. 7) zugrunde gelegt und damit sowohl der Kontext-, der Input-, der Produkt- und der Prozessperspektive Rechnung getragen. Bei der Betrachtung der Prozessperspektive wurde dabei zum einen der Projektprozess (Meeting-Evaluationen, Zwischenergebnis-Evaluationen usw.), zum anderen der Erkenntnisprozess formativ untersucht. Der Erkenntnisprozess wurde formativ nachgezeichnet. Dies geschah durch die didaktische Untersuchung eines *Blended-Learning*-Kurses und seiner Lernergebnisse unter Berücksichtigung der App-Nutzung. Zudem wurden auch die Angebote im Rahmen der *Face-to-Face*-Kurselemente einbezogen. Des Weiteren konnte eine summative Evaluation mittels Lernenden- und Lehrkräftebefragungen zum Kursabschluss das Gesamtbild vervollständigen. Hierbei kam mit dem Zusammenspiel von Interviews und Fragebogenerhebungen eine Kombination aus qualitativer und quantitativer Untersuchung zum Einsatz, die im Rahmen einer Triangulation ergänzend den Forschungsgegenstand beschrieb und die didaktischen Strukturen reflektierte. Dabei kam zudem länderübergreifend ein vergleichender Ansatz zum Tragen, der Gruppen und Vergleichsgruppen kriterienorientiert untersuchte. Um eine unabhängige Außenmeinung damit abgleichen zu können, wurde zudem eine externe Evaluation im Rahmen quantitativer Forschung durchgeführt. Die Verschränkung der beruflichen Spezifika drückt sich sowohl inhaltlich in der „AGnovel“-App mit ihren diversen Fachanknüpfungspunkten als auch in der Untersuchung von situativen Ansätzen im Lehr-/Lernprozess aus. Kontextuell sind hier die Ansprüche der Gesellschaften und Bildungssysteme der Partnerländer (Deutschland, Großbritannien, Irland, Italien, Litauen, Spanien und Zypern) zu berücksichtigen. Aufgrund der guten Forschungsergebnisse wird seit Ende 2019 eine weitere Prototypengestaltung mit Blick auf das Sprachenlernen im Grundschulbereich sowie für das Sprachenlernen für Migrantinnen und Migranten angestoßen.

Im Projekt „MATH“ (vgl. BEUTNER/PECHUEL 2019) wird empirisch der Einsatz von *Serious Games* sowie deren Nutzen und Akzeptanz untersucht. Die Ausrichtung erfolgt im Bereich der vorberuflichen schulischen Bildung und Ausbildungsvorbereitung. Dabei kommen ebenfalls mehrere Perspektiven zum Tragen: Auf der Mikroebene geht es um die Erforschung des didaktischen Einsatzes von *Serious Games* im Unterricht, auf der Mesoebene um die Frage, wie *Serious Games* organisatorisch in die Bildungsgänge einer Schule integriert und abgestimmt werden können. Auch hier kommt eine Kombination aus qualitativen und quantitativen Methoden zum Einsatz. Insbesondere die Nutzerfreundlichkeit wird im Rahmen einer *Usability*-Studie untersucht, die methodisch das modifizierte *Technology-Acceptance-Modell* (TAM) anwendet und dies um Interviews mit Lehrkräften, Schulverantwortlichen und Lernenden ergänzt. Untersucht werden die Nutzung des Prototyps und die Implementation dieses *Serious Games*. Dabei werden an jedem Untersuchungsstandort in den vier Ländern, in denen kooperativ die „MATH“-App entwickelt und der *Serious-Games*-Ansatz erforscht wurde, die schulischen Situationen und die curricular-didaktischen Einbettungen ermittelt. Partnerländer im Rahmen des „MATH“-Projektes sind Deutschland, Bulgarien, Spanien und

die Türkei. Dabei findet zum einen jeweils eine Lehrplananalyse statt, um passende organisatorische Anknüpfungsmöglichkeiten für die App auszumachen. Zum anderen werden die Auswirkungen auf Lehr-/Lernprozesse untersucht. Die Rahmenbedingungen des jeweiligen Bildungssystems des Landes, in dem „MATH“ zum Einsatz kommt, bilden jeweils den Kontext der Umsetzung und reichen damit von stark bis hin zu schwach staatlich regulierten Systemen. Die Tests der Prototypen zeigen derzeit eine gute Einbindungsmöglichkeit in den Unterricht, sofern in den Präsenzphasen tatsächlich auf die App und die Fragen der Lernenden Bezug genommen wird und die Theorie nicht losgelöst von der Umsetzung im *Serious Game* erfolgt. Diese Forschungsergebnisse sind einer der Gründe, warum als Rückwirkung auf den Bereich der Entwicklung auch *Bootcamps* an den Bergen im Spiel platziert wurden, die zusätzliche Lernhilfen beinhalten und den Lernenden eine ortsunabhängige Erläuterung der mathematischen Hintergründe situativ eingebunden ermöglichen.

In „NetEnquiry“ (vgl. BEUTNER 2016) wurde der Fokus auf die Ausbildung und Weiterbildung im Bankenbereich gelegt, wobei der Weiterbildungsbereich besonders betont wird. Während der Weiterbildungsbereich durch die Einbindung von Bankbetrieben ins Projekt aufgenommen wird, ist der Ausbildungsbereich sowohl betrieblich als auch berufsschulisch integriert. Der Fokus liegt auf der planerischen curricularen Ebene, sprich der Mesoebene, die Implementation erfolgt jedoch auf einer Kurs- bzw. unterrichtlichen Ebene. Der Schwerpunkt lag dabei auf der Untersuchung von Inputs, Prozessen und Outputs sowie deren Zusammenspiel und wurde multiperspektivisch über einen Methodenmix aus Fragebogenmethode, Interviewstudien als Einzel- und Gruppeninterviews sowie durch eine Toolanalyse abgebildet. Bei der Toolanalyse kam quantitativ auch hier eine modifizierte TAM-Betrachtung zur *Usability*-Untersuchung zum Einsatz, die aber durch ein *Szenario-Testing* qualitativer Art mit Blick auf die Bereiche technische Umsetzung, didaktische Umsetzung, Vollständigkeit, Richtigkeit, Prozessabbildung und Szenario-Gestaltung ergänzt wurde. Die Einzelinterviews wurden mit Lehrkräften, Entwicklern und Entwicklerinnen sowie Bankvertretern und Bankvertreterinnen umgesetzt und richteten den Blick vorrangig auf Nutzen, Herausforderungen, didaktische Einbettung, Potenziale und Szenario-Gestaltung. Demgegenüber fokussierten die Gruppeninterviews, die ebenfalls mit den Zielgruppen der Lehrkräfte, Entwicklern und Entwicklerinnen sowie Bankvertretern und Bankvertreterinnen durchgeführt wurden, Fragen der Authentizität, des Gesamtnutzens, des didaktischen Designs sowie des Transfers. Zum Transfer wurden zudem speziell die Berufskollegs angesprochen, da sich hier ein bildungsgangübergreifender Transfer ggf. speziell innerhalb von Institutionen, aber auch über Institutionengrenzen hinweg anbietet. Im quantitativen Teil kamen Fragebögen zur Akzeptanz von *Mobile Learning*, zu Zukunftsperspektiven des Berufes und der Digitalisierungstendenzen sowie zur Dozentenschulung und Fragebögen für Lernende zum Einsatz, die Zufriedenheit, didaktische Einbindungen, technisches und didaktisches Design abbilden.

Die durchweg guten und sehr positiven Ergebnisse seien hier anhand der Akzeptanzanalyse im Rahmen der *Usability*-Studie aufgezeigt. Auf einer Skala von 1 bis 6 jeweils ge-

messen über einer Operationalisierung mittels sechs Items liegt der wahrgenommene Bedienkomfort (PEOU – *Perceived Ease of Use*) bei einem Wert von 5,29, die Einstellung zur Nutzung (ATTITUDE – *Attitude towards Using*) bei 5,66, die wahrgenommene Nützlichkeit (PU – *Perceived Usefulness*) bei 5,59 und die Nutzungsabsicht (ITU – *Intention to Use*) ebenfalls bei 5,59. Ergänzt wurde dies durch Profilvergleichungen zur Beschreibung von *Mobile Learning*, um eine ganzheitliche Annäherung an diesen zu untersuchenden Themenbereich zu erlangen. Die Kontextualisierung erfolgt bei diesem innerdeutschen Projekt über die Berücksichtigung der unterschiedlichen Rahmenbedingungen der Länder: Diese schlagen sich unmittelbar in den berufsspezifischen Besonderheiten nieder, z. B. in den unterschiedlichen länderspezifischen curricularen Vorgaben im Rahmen der Lehrpläne, die einer einzigen Ausbildungsordnung gegenüberstehen, und in den organisatorischen Besonderheiten der verschiedenen Bankenformen, die alle im Projekt repräsentiert waren (Privatbanken-, Genossenschaftsbanken- und Sparkassensystem).

5 Zusammenführung und Ausblick

Insgesamt kann festgehalten werden, dass mit dem Dreischalenmodell der systematisch-analytischen Berufsbildungsevaluation ein speziell auf die Bedürfnisse der Berufsbildung ausgerichtetes Orientierungsraster vorgelegt wird, dass die Anwenderinnen und Anwender nicht zu spezifischen Vorgehensweisen drängt, sondern Bereiche aufzeigt, in denen Entscheidungen zu treffen sind und zu denen im Rahmen der Planung, Gestaltung und Ausformung von Evaluationen Stellung genommen werden sollte. Herausforderungen für die künftige Weiterentwicklung sind etwa Fragen der grundlegenden Normen und die Frage, ob und wie solche spezifischer ausgeformt werden sollten, wenn es um die Anwendung des Modells geht. Deutlich wurde jedoch bereits bei der Breite der bisherigen Anwendungen, dass sowohl unterschiedlichen Forschungsparadigmata wie auch unterschiedlichen Perspektiven Rechnung getragen werden kann. Insgesamt bietet das Modell damit einen spezifischeren Zugang als rein allgemeine Evaluationsmodelle.

In der Umsetzung durch Dritte kommt es bisweilen zu Redundanzen bei der Betrachtung des Kontextes (Schale 3) und den berufsspezifischen Besonderheiten (Schale 2). Dies liegt jedoch in der Regel an der Deutung von allen Rahmenbedingungen als Kontexte. Eine solche Deutung ist vom Modell in dieser Form nicht intendiert. Zudem werden bisweilen die Berufsbildungsspezifika der Schale 2 mit den Kernorientierungen im Zentrum des Modells vermischt. Hier ist aber bewusst Schale 2 nicht als Schale 1 gewählt worden, um eine Vermischung von genereller Ausrichtung von Bereichen und Ebenen (Kern) mit den Besonderheiten, die in der Berufsbildung vorliegen und in der Regel einer viel dezidierteren Betrachtung bedürfen, zu vermeiden. Es soll zunächst eine allgemeine Orientierung (Kern) erfolgen, dann daraufhin eine methodisch adäquate Annäherung, bevor in den Überlegungen eine Einbindung berufsspezifischer Besonderheiten Eingang finden soll. Abschließend wird eine Berücksichtigung des Kontextes mit aufgenommen, damit sich der Blickwinkel zum Ende

wieder weitet. Diese Abfolge hat sich inzwischen bei mehr als 50 Umsetzungen bewährt und wurde in der Modellgenese mehrfach umstrukturiert, bis sich die aktuelle Version ergab, die nach bisherigen Untersuchungen und Erfahrungen zu den geringsten Redundanzen bei einer klaren Orientierung führt. Hier wird die Zukunft zeigen, wie die ausstehenden Fragen zu ggf. orientierungsstiftenden Normen sinnvoll einzubinden sind.

Literatur

- ACHTENHAGEN, Frank: Zur Evaluation komplexer Lehr-Lernarrangements als neue Formen des Lehrens und Lernens in beruflichen Schulen. In: GONON, Philipp (Hrsg.): Evaluation in der Berufsbildung. 2. Aufl. Aarau, Frankfurt am Main 1995
- BEUTNER, Marc: Evaluationsmodelle und der Kontext der beruflichen Bildung. In: Kölner Zeitschrift für Wirtschaft und Pädagogik 22 (2007) 43, S. 79–101
- BEUTNER, Marc: Evaluation und beruflichen Bildung. Festvortrag im Rahmen der Festveranstaltung „60 Jahre Wirtschaft und Erziehung“, Leipzig, am 06.11.2008. In: Wirtschaft und Erziehung (2009) 11/12
- BEUTNER, Marc: NetEnquiry. Innovative Ansätze zum Serious Mobile Learning für Aus- und Weiterbildung. Köln 2016
- BEUTNER, Marc: Berufsbildungsevaluation. Ein Lehrbuch für Berufs- und Wirtschaftspädagogen, Studierende des Lehramts an berufsbildenden Schulen sowie Theorie und Praxis. 2. Aufl. Köln 2018
- BEUTNER, Marc: Die DeGEval-Standards in der beruflichen Bildung. In: HENSE, Jan Ulrich u. a. (Hrsg.): Evaluation: Standards in unterschiedlichen Handlungsfeldern. Einheitliche Qualitätsansprüche trotz heterogener Praxis? Münster 2019, S. 121–136
- BEUTNER, Marc; PECHUEL, Rasmus: NetEnquiry. Modernes M-Learning im Bankenbereich. Vortrag im Rahmen des eLearning Summits am 12.11.2013 in München im Rahmen der eLearning Summit Tour 2013. München 2013
- BEUTNER, Marc; PECHUEL, Rasmus: App-based Interactive Language Learning – Pedagogy, Implementation and Evaluation with the AGnovel Tool. In: Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2015, S. 1408–1414 – URL: <http://www.editlib.org/p/152181/> (Zugriff: 14.01.2019)
- BEUTNER, Marc; PECHUEL, Rasmus: Mountains of Absolutely Terrifying Height (MATH) – Creating a Serious Games through Design-Based Research. SITE 2019. Society for information technology and teacher education. Las Vegas 2019
- BUSCHFELD, Detlef: Konditionen beruflicher Bildungsgänge. Theoretische Fundierung eines berufs- und wirtschaftspädagogischen Konzepts. Köln 2002
- DEGEVAL – GESELLSCHAFT FÜR EVALUATION: DeGEval – Gesellschaft für Evaluation. Standards für Evaluation. Erste Revision auf Basis der Fassung 2002. Verabschiedet durch die Mitgliederversammlung der DeGEval e. V. am 21. September 2016 – Langfassung – URL:

https://www.degeval.de/fileadmin/Publikationen/DeGEval_Standards_fuer_Evaluation_-_Erste_Revision__2016_.pdf (Zugriff: 14.01.2019)

- GEORG-SIMON-OHM BERUFSSKOLLEG (Hrsg.): Evaluationsportal. 2019 – URL: <http://www.gso-koeln.de/index.php/einrichtungen/238-evaluationsportal> (Zugriff: 14.01.2019)
- KRANERT, Thomas; KREMER, H.-Hugo: Innovationsarena 3i. Qualitätsmanagement in und von Entwicklungs- und Innovationsprozessen. Paderborn 2017 – URL: https://www.uni-paderborn.de/fileadmin/cevet/EX6_QM_in_und_von_Entwicklungs-_und_Innovationsprozessen.pdf (Zugriff: 14.01.2019)
- KROMREY, Helmut: Empirische Sozialforschung. Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung. Stuttgart 2006
- MAYRING, Philipp: Kombination und Integration qualitativer und quantitativer Analyse [31 Absätze]. In: Forum Qualitative Sozialforschung / Forum Qualitative Social Research 2 (2001) 1, Art. 6 – URL: <http://nbn-resolving.de/urn:nbn:de:0114-fqs010162> (Zugriff: 14.01.2019)
- PATTON, Michael Quinn: Developmental evaluation: applying complexity concepts to enhance innovation and use. New York 2010
- POPPER, Karl R.: Falsifizierbarkeit, zwei Bedeutungen von. In: SEIFFERT, Helmut; RADNITZKY, Gerard (Hrsg.): Handlexikon zur Wissenschaftstheorie. München 1989, S. 82–85
- POPPER, Karl R.: Logik der Forschung. Hrsg. von Herbert Keuth, 11. Aufl. Tübingen 2005
- SCRIVEN, Michael: Die Methodologie der Evaluation. In: WULF, Christoph (Hrsg.): Evaluation. München 1972, S. 60–91
- SCRIVEN, Michael: Evaluation Thesaurus. 4. Aufl. Thousand Oaks 1991
- SCRIVEN, Michael: Conceptual Revolutions in Evaluation: Past, Present, and Future. In: ALKIN, Marvin C. (Hrsg.): Evaluation Roots. A Wider Perspective of Theorists' Views and Influences. 2. Aufl. Los Angeles 2004a, S. 167–179
- SCRIVEN, Michael: The Fiefdom Problem. In: Journal of MultiDisciplinary Evaluation 1 (2004b) 1, S. 11–18 – URL: http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/143/15 (Zugriff: 14.01.2019)
- SCRIVEN, Michael: The logic of evaluation. Claremont Graduate University. 03.07.2007 – URL: <http://homepages.wmich.edu/~mscriven> (Zugriff: 14.01.2019)
- SLOANE, Peter F. E.: Schulnahe Curriculumentwicklung. In: BWP@ Berufs- und Wirtschaftspädagogik – online (2003) 4 – URL: http://www.bwpat.de/ausgabe4/sloane_bwpat4.shtml (Zugriff: 14.01.2019)
- STECKLER, Allan u. a.: Toward integrating qualitative and quantitative methods: An introduction. In: Health Education Quarterly (1992) 19, S. 1–8
- STUFFLEBEAM, Daniel L.: The CIPP model for evaluation – an update – a review of the model's development – a checklist to guide implementation. Portland 2003

Wolfgang Meyer

► Wirkungsverläufe messen: eine kleine Sinfonie der Berufsbildungszusammenarbeit

Viele Wissenschaftlerinnen und Wissenschaftler sehen experimentelle Designs als den einzig sinnvollen Weg zur Messung von Wirkungen an. Dieser Beitrag argumentiert in eine andere Richtung und zieht den Vergleich zu einem sinfonischen Werk. Erörtert wird in einem ersten Schritt die Frage des Wirkungsverlaufs, d. h. die Entwicklung von Effekten in Abhängigkeit von unterschiedlichen Instrumenten und deren spezifischen zeitlichen Perspektiven, die sich einer vergleichenden Pauschalbeurteilung entziehen. In einem zweiten Schritt wird auf das Zusammenspiel verschiedener Instrumente und die daraus folgende Komplexität gemeinsam erzielter Wirkungen eingegangen. Experimente sind von Natur aus reduktionistisch und zur Bewertung von Systemwirkungen wenig geeignet. Abschließend wird anhand eines Praxisbeispiels die Langfristigkeit der Zusammenarbeit hervorgehoben und die daraus folgende Aufgabe, die Instrumente entsprechend den sich stetig verändernden Rahmenbedingungen und Anforderungen dauerhaft vorausschauend möglichst effektiv zu dirigieren. Was gestern gewirkt hat, muss nicht morgen wirken – und umgekehrt. Berufsbildungszusammenarbeit stellt deshalb Anforderungen, welche eher durch begleitende und längsschnittorientierte Verfahren im Methodenmix als durch umfassende punktuelle und primär quantitativ gemessene Experimente erfüllt werden können. Auch hierfür wird ein Praxisbeispiel angeführt.

1 Präludium in E – eine etwas andere Einleitung

In vielen Fachgebieten werden experimentelle Untersuchungsdesigns (üblicherweise nach der englischen Abkürzung von „randomised controlled trials“ als RCTs bezeichnet) als „Goldstandard“ der Wirkungsmessung angesehen (z. B. in der Medizin vgl. kritisch KAPTCHUK 2001). Auch bei der Längsschnittbetrachtung von Wirkungen in der beruflichen Bildung wird gelegentlich der Anspruch vertreten, dass experimentelle Verfahren am besten hierzu geeignet sind (vgl. z. B. PRESCOTT-CLEMENTS u. a. 2008). Dieser Beitrag argumentiert genau in die entgegengesetzte Richtung: Experimentelle Designs sind aufgrund ihrer methodischen Ausrichtung und der Besonderheiten der Berufsbildungszusammenarbeit zur Wir-

kungsmessung wenig geeignet; es sind vielmehr andere, längsschnittorientierte Methoden notwendig.

Die fehlenden Kenntnisse zu den Wirkungsverläufen in der Beruflichen Bildung und deren Kontextabhängigkeit stellen dabei in diesem Beitrag den ersten Beleg für diese These dar. Der nächste Abschnitt (Kapitel 2) macht deutlich, dass experimentelle Designs zur Erfassung dieser Spezifika wenig geeignet sind und unter bestimmten Umständen sogar schädlich sein können.

Der darauffolgende Teil (Kapitel 3) fokussiert eine zweite wichtige Komponente der Berufsbildungszusammenarbeit, die einer Prüfung der Wirkungszusammenhänge mittels experimenteller Designs entgegensteht. Zumeist handelt es sich nämlich um eine langfristige Kooperation zwischen unterschiedlichen Partnern, die in einem Netz unterschiedlicher Aktivitäten auf verschiedenen Interventionsebenen miteinander verbunden sind und die sich nicht durch die reduktionistische Logik eines Experiments erfassen lassen.

Zum Abschluss dieses Beitrags (Kapitel 4) zeigt ein Beispiel, auf welche Art und Weise methodisch an die Problematik einer langfristigen Wirkungsmessung in komplexen Wirkungsgeflechten herangegangen werden kann. Damit wird sicher nicht der Anspruch einer kompletten Erfassung und Prüfung der Wirkungsmechanismen erhoben, aber zumindest der Weg für eine andere, der Realität der Berufsbildungszusammenarbeit angemessenere methodische Herangehensweise aufgezeigt.

Dies geschieht in Analogie zur Evaluation sinfonischer Musik, deren Wirkungen sich in ähnlicher Weise wie die Berufsbildungszusammenarbeit einer Evaluation mittels experimenteller Vorgehensweise entziehen. Entstanden ist so eine kleine Sinfonie der Berufsbildungszusammenarbeit, die ähnlich wie in der Musik das harmonische Zusammenspiel mehrphasiger Inventionen in den Fokus wissenschaftlicher Wirkungsforschung stellt.

2 Der Goldstandard und die Berufsbildungsrealität mit ihren Wirkungsverläufen

Selbstverständlich lässt sich auch eine Sinfonie experimentell untersuchen: Am Eingang des Konzerthauses werden die Besucher/-innen randomisiert zwei verschiedenen Konzertsälen zugewiesen, in einem der beiden Räume wird die Sinfonie aufgeführt und im anderen nicht, am Ende wird schließlich der Pegel des Applauses zur Wirkungsbeurteilung gemessen. In der Tat ist den Künstlerinnen und Künstlern eine solche „Evaluation“ des Erfolgs ihres Auftritts mittels Applauses sehr wichtig und trotzdem wird schnell deutlich, dass weder Messverfahren noch methodische Konzeption dem Anspruch einer Wirkungserfassung des sinfonischen Werks gerecht werden können.

Da ist zunächst der Sachverhalt, dass eine simple Pauschalbeurteilung die komplexe „Interventionslogik“ der Musik verfehlt. Die experimentelle Vorgehensweise folgt einem reduktionistischen Denken: Es geht darum, eine möglichst einfache Intervention unter weitgehender Konstanthaltung der Randbedingungen im Vergleich zweier gleicher Gruppen zu

untersuchen, die dieser Intervention ausgesetzt sind oder nicht. Wenn eine Veränderung in der Interventionsgruppe festgestellt wird, die sich in der Kontrollgruppe nicht ergeben hat, so lässt sich aufgrund der Konstanzhaltung der Rahmenbedingungen sagen, dass diese ausschließlich auf die Intervention zurückzuführen ist. Setzt sich die Intervention allerdings aus verschiedenen Teilelementen zusammen, so bleibt unklar, welches davon Effekte auslöst oder ob sie sich eventuell gegenseitig neutralisieren. Das Experiment beantwortet also nur die Frage, ob es eine Wirkung der untersuchten Merkmalskombination gegeben hat, nicht jedoch die Fragen nach dem Warum, dem Wie oder danach, wie lang Wirkungen über die Zeit erhalten bleiben (vgl. als Überblick zum Einsatz von Experimenten in den Sozialwissenschaften DENZIN 2009).

Gerade der letzte Aspekt wird von den Protagonisten und Protagonistinnen experimenteller Designs häufig übersehen: Wirkungen entfalten sich keineswegs immer gleich und bleiben ebenfalls nicht immer in derselben Form weiter bestehen. Implizit geht das Experiment davon aus, dass eine durch die Intervention ausgelöste Veränderung zum Zeitpunkt der Nachher-Messung erfasst werden kann. Dies funktioniert sicherlich immer dann, wenn durch die Intervention ein dauerhafter Statuswechsel erzielt wird (z. B. ein Berufsabschluss erreicht wurde oder nicht).

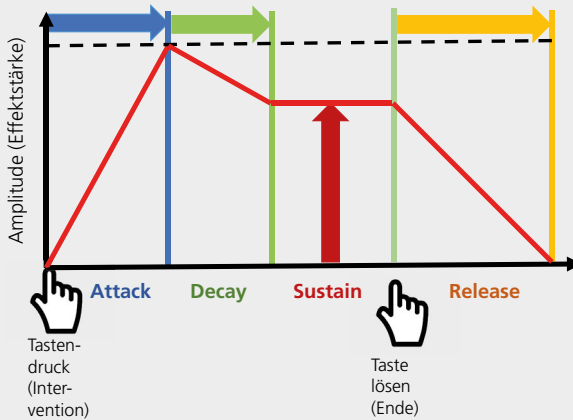
Leider sind solche Effekte aber bei Bildungsinterventionen eher die Ausnahme denn die Regel: Auch wenn sich viele Lehrerinnen und Lehrer wünschen würden, dass sie in den Köpfen ihrer Schülerinnen und Schüler nur einen Schalter umzulegen bräuchten, verlaufen Lernprozesse in Wahrheit komplizierter und individuell höchst unterschiedlich.¹ Ähnliches gilt übrigens auch für die Musik: Die einzelnen Instrumente unterscheiden sich hinsichtlich der Tonerzeugung und der dadurch erzielten Wirkungen bei den Zuhörerinnen und Zuhörern.

Bei modernen Instrumenten wie z. B. dem Synthesizer lassen sich diese Effekte manipulieren und gezielt einsetzen. In der Musiksynthese wird dabei besonders das ADSR-Hüllkurvenmodell verwendet (vgl. SCHIFFNER 2003, S. 50f., zur Umsetzung in der analogen Musiksynthese PINCH/TROCCO 2009). In diesem allgemeinen Modell werden vier verschiedene Elemente der Signalentwicklung differenziert: *Attack*, *Decay*, *Sustain* und *Release* (siehe Abbildung 1).

1 Dies gilt im Besonderen für die Berufsbildungszusammenarbeit, in der z. B. die Wege aus der Armut durch berufliche Bildung als komplexe Wirkungsverläufe zu beschreiben sind (vgl. z. B. LANGTHALER 2013, S. 3).

Abbildung 1: Das ADSR-Hüllkurvenmodell

ADSR-Hüllkurvenmodell



Quelle: eigene Darstellung in Anlehnung an <https://www.wikiwand.com/de/ADSR> (Zugriff: 08.01.2021)

Die *Attack*-Phase beschreibt den Zeitraum zwischen der Intervention (z.B. dem Tastendruck) und dem Moment, an dem das Signal (der Ton) die maximale Ausprägung (Lautstärke) erlangt. Nach diesem Zeitpunkt startet eine zweite Phase, der *Decay*, bei dem der Ton auf ein dauerhaftes Niveau (den *Sustain*) bei weiterhin gedrückter Taste absinkt. Mit dem Lösen der Taste startet schließlich eine dritte Phase des Ausklingens des Tons (*Release*). In ähnlicher Weise lässt sich auch die Entwicklung einer Berufsbildungsmaßnahme beschreiben: Ein neues Schulungsverfahren entfaltet erst allmählich seine Wirkungen (*Attack*), kann dann aber aufgrund von Gewöhnungseffekten wieder an Wirksamkeit verlieren (*Decay*) und sich auf einem bestimmten Niveau stabilisieren (*Sustain*). Nach Ende einer Intervention geht häufig der Effekt weiter zurück (*Release*) – um sich im Unterschied zur Musik langfristig auf niedrigerem Niveau zu stabilisieren (sozusagen eine zweite *Sustain*-Phase).

Im Experiment sind diese Unterschiede im Wirkungsverlauf nicht zu erfassen, weil es – zumindest in einer klassischen Vorgehensweise – an Messzeitpunkten fehlt. Es bedarf einer kontinuierlichen Wirkungsbeobachtung, welche über die Zeit vergleichbare Messungen durchführt – und nicht „nur“ eines Vorher-nachher-Vergleichs verschiedener Gruppen. Zusätzlich verweist das Beispiel auf ein zentrales Problem bei der Durchführung von Experimenten: die Terminierung der Nachher-Messung. Während die zeitliche Festlegung der Vorher-Messung mehr ein praktisches Problem darstellt – viele Interventionen zeichnen sich durch das Fehlen einer Baseline-Studie aus – hängt das Ergebnis einer experimentellen Wirkungsmessung von der Bestimmung des zweiten Messpunkts und dem theoretisch zu vermutenden Wirkungsverlauf ab. Eine Vorher-nachher-Messung im Sinne von vor und

nach dem Tastendruck führt in der Musik zu unterschiedlichen Ergebnissen bezüglich der Wirkung (Lautstärke des Tons), je nachdem ob noch in der *Attack*-, erst in der *Decay*- oder schon in der *Release*-Phase gemessen wird. Gleiches gilt für Interventionsmaßnahmen in der Berufsbildung, die ebenfalls nur selten zu konstant gleichen Wirkungen führen.

Noch ein weiterer Punkt spricht gegen den „Goldstandard“ und für eine kontinuierliche Wirkungsbeobachtung in der Berufsbildungszusammenarbeit. Es geht in der Regel gar nicht primär um die Frage, ob eine Intervention wirkt oder nicht. Zumeist werden im internationalen Kontext Maßnahmen eingesetzt, die sich schon lange im deutschen Kontext bewährt haben und von deren generellen Wirksamkeit deshalb ausgegangen werden kann. Auch der Musiker oder die Musikerin stellt bei einem Instrument nicht infrage, dass damit prinzipiell ein Ton zu erzeugen ist – wichtig ist aber z. B. die Frage, ob sich das Instrument aufgrund der Kontextbedingungen verstimmt hat. Genauso stellt in der Berufsbildungszusammenarbeit die Variation der Kontextbedingungen und deren Einfluss auf die erwarteten (weil durch Erfahrungen bekannten) Wirkungen von Maßnahmen den eigentlichen Untersuchungsgegenstand dar. Experimente helfen hier nicht oder nur bedingt weiter.

3 Wie die Musik spielt – langfristige Kooperation statt einmaliger Intervention

Mit dem Stimmen der Instrumente ist allerdings weder in der Musik noch in der Berufsbildungszusammenarbeit die Wirkungsbeobachtung abgeschlossen. Eine Sinfonie besteht nicht aus einem Ton eines Instrumentes und Berufsbildungszusammenarbeit lässt sich nicht aufgrund der Wirksamkeit einer bestimmten Maßnahme zu einem bestimmten Zeitpunkt beurteilen. Der von Johann Sebastian Bach in die Musiklehre eingeführte Begriff der Sinfonie kommt aus dem Griechischen und bedeutet „zusammenklingend“, „harmonisch“ – und genau darum geht es in der Musik (nicht nur in der Sinfonie). Mehrere Instrumente spielen unterschiedliche Dinge, die sich erst im Zusammenspiel zu einem musikalischen Werk vereinen und dadurch gemeinsam eine Wirkung entfalten. Ein „falscher Ton“ oder eine abweichende Rhythmik zerstören das Gesamtkunstwerk – das richtige Spiel der einzelnen Instrumente (und die damit erzielte Wirkung) stellt eine notwendige, aber keineswegs eine hinreichende Bedingung für den Erfolg der Sinfonie dar.

Auch hier liegt die Parallele zur Berufszusammenarbeit auf der Hand: Internationale Zusammenarbeit beschränkt sich nicht auf einzelne Maßnahmen, sondern basiert auf einem in der Planungsphase entwickelten, in der Regel sehr komplexen Zusammenspiel unterschiedlicher Instrumente (zumeist auf verschiedenen Interventionsebenen), deren Erfolg nicht nur von der Wirksamkeit des einzelnen Instruments, sondern vor allem auch von deren Zusammenspiel abhängt (als Überblick zur Entwicklung der Berufsbildungszusammenarbeit und ihrer Instrumente vgl. HERTMANN 2019). Kleinste Störungen oder Verzögerungen können die Wirksamkeit be- oder sogar verhindern. Eine noch so geschickt durchgeführte experimentelle Studie, die am Ende der Zusammenarbeit den Erfolg bzw. Misserfolg wissenschaftlich

prüft, hilft während der Durchführung der Berufsbildungszusammenarbeit nicht. Sie trägt auch wenig zu einer Verbesserung der nächsten Kooperation bei, die in einem anderen Kontext unter abweichenden Bedingungen durchgeführt wird. In beiden Fällen geht es nicht um die Feststellung einer Gesamtwirkung, sondern um eine Analyse der Umstände, die Fehler als Folge hatten (oder haben könnten) und deren Früherkennung für das Gelingen des Zusammenspiels essenziell ist. Die Tatsache, dass eine Sinfonie insgesamt schlecht geklungen hat, hilft dem Dirigenten oder der Dirigentin nicht weiter – es geht um die Identifikation von Fehlerquellen und Maßnahmen, die zu einer Verbesserung führen.

Die Analogie zur Musik kommt aber an ihre Grenzen, weil es sich bei der Berufsbildung generell um eine kontinuierliche Tätigkeit handelt und auch die Berufsbildungszusammenarbeit nicht unbedingt zu einem Ende kommen muss, sondern sich auf verschiedene, parallel verlaufende und im Wandel befindliche Prozesse bezieht (vgl. z. B. die Beschreibung der Aufgaben von Aus- und Weiterbildung durch GESSLER/Kühn 2019, S. 36). Berufe verändern sich in Abhängigkeit von der sozioökonomischen und technischen Entwicklung über die Zeit, und bei manchen Berufsausbildungen gibt es sicher auch ein „natürliches“ Ende (dies wird gegenwärtig insbesondere bezüglich der Folgen der vierten industriellen Revolution diskutiert, vgl. WINDELBAND 2014 oder die Szenario-Rechnungen des IAB: WOLTER u. a. 2015). Ob dies aber Folgen für das Berufsbildungssystem oder gar die Berufsbildung insgesamt hat, ist zumindest strittig und hinsichtlich seines Ausmaßes keineswegs eindeutig zu beantworten (vgl. für entsprechende Annahmen BOSCH 2014). Bisher hat das deutsche duale Berufsbildungssystem jedenfalls noch alle Kassandrarufer überlebt und sich als äußerst robust und anpassungsfähig erwiesen. Dies bedeutet jedoch mitnichten, dass alle die berufliche Bildung prägenden Lehr-/Lern-, Vermittlungs-, betrieblichen Produktions- und gesellschaftlichen Kontextprozesse über die Zeit stabil geblieben sind. Dies trifft ebenso auf die unterschiedlichen Implementationsprozesse beim Transfer beruflicher Ausbildungen in bestehende schulische bzw. betriebliche Kontexte zu (vgl. die Ergebnisse der Fallstudie von GESSLER 2017).

Die neuere Entwicklung in der Berufsbildungszusammenarbeit (und nicht nur dort) geht dahin, Berufsbildung nicht als eigenständige, abgeschlossene Phase zu begreifen, sondern als Teil des „lebenslangen Lernens“ und sie damit enger als bisher mit anderen Bildungsphasen zu verknüpfen (vgl. hierzu ARNOLD/FEDER 2005, S. 439ff.). Hinzu kommt die Tatsache, dass insbesondere Nichtregierungsorganisationen sehr partnerschaftlich orientiert sind und für sie generell die langfristige Zusammenarbeit mit denselben Partnern zu deren *Empowerment* im Vordergrund steht (vgl. LENZEN 2001, S. 8ff.).

Die Berufsbildungszusammenarbeit der Hanns-Seidel-Stiftung (HSS) in der Volksrepublik China ist ein gutes Beispiel dafür, wie im Bereich der Nichtregierungsorganisationen eine solche langfristige Kooperation gestaltet werden kann (vgl. ausführlich STOCKMANN/MEYER 2017, insbesondere S. 74ff.). Die Ursprünge reichen bis zur vorsichtigen Annäherung des damaligen CSU-Vorsitzenden Franz Josef Strauß und der chinesischen Führung unter Mao Zedong Mitte der 1970er-Jahre zurück, die schließlich 1983 zu einer bis zum heutigen Tag anhaltenden Zusammenarbeit der HSS mit der Berufsschule für Bauberufe in Nanjing

führte. Diese war für beide Seiten eine ungemaine Herausforderung: Die HSS als politische Stiftung war zu diesem Zeitpunkt in der Berufsbildungszusammenarbeit noch unerfahren und im chinesischen Berufsbildungssystem fehlte es so ziemlich an allem: ausgebildeten Lehrkräften, fachpraktischen Curricula, Ausbildungsmaterialien und für die Berufsbildung angemessenen Räumlichkeiten.

Sukzessive – mit langem Atem und bescheidenem Mitteleinsatz (insbesondere im Vergleich zur staatlichen Berufsbildungszusammenarbeit, vgl. STOCKMANN u. a. 2000, S. 39ff.) – hat die HSS in den 1990er-Jahren ihr Engagement in der Berufsbildungszusammenarbeit in der Volksrepublik China ausgeweitet. Wichtige Meilensteine waren der Aufbau eines Zentrums für Lehrerfortbildung in Shanghai, eines Management-Training-Centers in Hangzhou und eines berufspädagogischen Fortbildungszentrums in Qingzhou, die Kooperation mit weiteren Berufsbildungszentren in Shiyan, Shanghai, Wuhan, Weifang, Nanjing, Pingdu und Jiuquan mit Ausbildungsgängen vorrangig in den Bereichen Metall, Elektro und Landwirtschaft sowie die Anbahnung zahlreicher Kooperationen und Austauschprogramme mit (vorrangig bayrischen) Berufsbildungseinrichtungen und berufspädagogischen Hochschulen. Auch wenn Schwerpunkte verlagert und dadurch Kooperationen mit einzelnen Einrichtungen phasenweise stärker oder schwächer im Fokus der Arbeit der HSS standen, hat die Stiftung niemals den Kontakt vollständig abgebrochen und damit eines ihrer Projekte endgültig beendet. Es ist im Gegenteil Philosophie der Zusammenarbeit, die Einrichtungen in China als Teil der HSS-Anstrengungen zu verstehen und sie aktiv in die eigene Arbeit einzubinden. Über die vielen Jahre sind so ein gemeinsames Verständnis der Berufsbildung und eine verlässliche Kooperation gleichrangiger Partner entstanden, auf denen die HSS immer neue Aktivitäten aufsetzt und alte kontinuierlich weiterentwickelt. Zwar haben sich über die Zeit der Mittel- und Personaleinsatz der Stiftung verringert, der Umfang und die Intensität der Arbeit vor Ort haben sich dabei aber eher vergrößert.

Entstanden ist ein Netzwerk unterschiedlicher Einrichtungen, das über mehrere chinesischen Provinzen hinweg zusammenarbeitet und in einem komplizierten Wirkungsgeflecht miteinander verbunden ist. Das dauerhafte Zusammenspiel der einzelnen Elemente und ihrer verschiedenen Maßnahmen stellt die Stärke dieses Konzepts dar. Das größte Problem ist allerdings die kontinuierliche Wirkungsbeobachtung und -bewertung, die für eine systematische Weiterentwicklung dringend notwendig ist. Hier hat die HSS bisher noch keine überzeugende Lösung gefunden. Experimentelle Designs – so viel ist jedenfalls sicher – hätten in diesem Kontext kaum einen großen Nutzwert.

4 Methoden zur Wirkungsevaluation langfristiger Kooperationen

Die Bewertung der Wirksamkeit eines Vorhabens der Berufsbildungszusammenarbeit ist keine Aufgabe, die sich durch eine Ex-post-Evaluation mithilfe eines experimentellen Designs beantworten lässt. Die große Herausforderung besteht vielmehr darin, mit einem Blick voraus möglichst frühzeitig nahende Schwierigkeiten oder Fehlerquellen zu entdecken und

hierdurch die Wirksamkeit von Maßnahmen abzusichern. Ferner geht es darum, mit einem Blick zurück den tatsächlich geleisteten Beitrag an eingetretenen Veränderungen zu bilanzieren und aus den Grenzen der Wirksamkeit eigener Beiträge Konsequenzen für das weitere Handeln abzuleiten. Schließlich besteht die dritte Herausforderung darin, diese Erkenntnisse möglichst schnell und „*just in time*“ vor jeweils anstehenden Richtungsentscheidungen zu erlangen.

Für die Projektsteuerung ist eine präzise Bestimmung, ob überhaupt Wirkungen eingetreten sind, weniger wichtig als die Identifizierung von Barrieren und Handlungsmöglichkeiten. Die Tatsache, dass etwas in der Vergangenheit wirksam gewesen ist, hilft bei Entscheidungen über das zukünftige Handeln weiter – sofern gleichzeitig erkannt wird, warum die Maßnahmen wirksam waren und unter welchen Umständen sie es nicht sind. Ferner bedarf es Informationen über den Wirkungsverlauf, d. h., ob es z. B. zu Verzögerungen beim Eintritt von Wirkungen gekommen ist oder ob das erreichte Wirkungsniveau gehalten werden konnte. All diese Informationen können über experimentelle Designs – so aufwendig sie auch sein mögen – nicht gewonnen werden. Es bedarf einer Längsschnittbetrachtung und eines geschickten Zusammenspiels von Wirkungsmonitoring und Wirkungsevaluation.

Die kontinuierliche Beobachtung der Wirkungsentwicklung sowie die Identifizierung von Meilensteinen im Rahmen des Wirkungsverlaufs sind zunächst die Aufgabe des Wirkungsmonitorings. Die theoretischen und methodischen Anforderungen sind dabei nicht zu unterschätzen: Häufig wird in der Praxis primär die Machbarkeit und damit die leichte Umsetzung von Indikatoren in den Vordergrund gestellt – mit manchmal fragwürdigen Annahmen über die nicht beobachtbaren Zusammenhänge zwischen dem Messergebnis des Indikators und dem durch ihn abzubildenden Sachverhalt (vgl. MEYER 2016, S. 18f.). Wenn methodische Fragen diskutiert werden, so betrifft dies zumeist die Messgenauigkeit der verwendeten Indikatoren. Dieses sicherlich gerade in den Sozialwissenschaften wichtige Thema – letztlich sind die Einflüsse von Störgrößen größer als in den Naturwissenschaften und führen deshalb zu einer stärkeren fehlerbedingten Streuung der Messwerte – berücksichtigt aber selten die Tatsache, dass eine (aufwendige) Steigerung der Messpräzision konterkariert wird von einer unpräzisen Abbildung des eigentlich interessierenden Gegenstands. Es geht beim Wirkungsmonitoring nicht um eine möglichst genaue Messung der Indikatorenwerte, sondern um eine korrekte Abbildung der Wirkungen und ihrer Entwicklungen.

Unter anderem fehlt den Praktikerinnen und Praktikern der Berufsbildung dafür das benötigte Wissen über die Wirkungsverläufe von Maßnahmen und die vorhandenen Möglichkeiten, diese mit einfachen Mitteln weitgehend adäquat abzubilden. Hier ist die Berufsbildungsforschung gefordert, die zwar dank aufwendiger experimenteller Untersuchungen wissenschaftlich gut abgesicherte Informationen zur generellen Wirksamkeit von Maßnahmen bereitstellt, selten aber an der Beschreibung von Wirkungsverläufen sowie deren Abbildung mit geeigneten Indikatoren arbeitet (Ausnahmen betreffen z. B. die stufenweise Vermittlung berufsmotorischer Fertigkeiten vgl. SCHELTEN 2009). Gerade im Kontext der Berufsbildungszusammenarbeit sollten primär Maßnahmen zum Einsatz kommen, über deren

Wirksamkeit verlässliche Informationen vorliegen und deren Wirkungsverlauf sich gut über handhabbare Indikatoren abbilden lässt.

Das beste Wirkungsmonitoring kann allerdings nur den gegenwärtigen Zustand abbilden und lediglich im begrenzten Maße Auskunft über zukünftige Entwicklungen geben. Die abgebildeten Zukunftserwartungen basieren auf theoretischen Annahmen über die Gleichförmigkeit von Wirkungsverläufen, die mit einer konkreten Maßnahme verbunden sind und unter bestimmten Rahmenbedingungen so auftreten.

Um nochmals die Analogie zur Musik zu verwenden: Wenn eine Taste gedrückt wird, geht der Musiker oder die Musikerin von einer immer gleichen Tonproduktion aus. Der Verlauf lässt sich mit der ADSR-Hüllkurve beschreiben. Er kann selbst bewusst oder unbewusst (z. B. durch die Stärke des Anschlags oder den Einsatz von Filtern am Synthesizer) Einfluss auf diesen Verlauf nehmen; es kann aber auch zu technischen Störungen aufgrund bestimmter Rahmenbedingungen kommen (z. B. führt Hitze zur Verstimmung von Instrumenten). Musiker/-innen achten deshalb penibel auf bestimmte Einflussfaktoren, die ihrem Instrument und dessen Tonproduktion schaden könnten. Nur so ist eine weitgehende Kontrolle ihres Spiels möglich.

Auf die Berufsbildungszusammenarbeit übertragen bedeutet dies, dass die Rahmenbedingungen der Wirksamkeit von Maßnahmen bekannt sein müssen. Sowohl die Rahmenbedingungen als auch die bereits erzielten Wirkungen (und die Wahrscheinlichkeit, in Zukunft Wirkungen entsprechend der Planung erzielen zu können) müssen kontinuierlich beobachtet werden, und im Falle von Abweichungen sind angemessene Gegenmaßnahmen zu treffen. Erschwerend kommt hinzu, dass die Aktivitäten in der Regel in Ländern stattfinden, die einem schnellen und schwer vorhersehbaren gesellschaftlichen Wandel unterliegen, der einerseits die berufliche Bildung beeinflussen kann, andererseits aber auch durch sie selbst beeinflusst wird – und nicht immer nur in die intendierte Richtung. Mit anderen Worten: Im Unterschied zur Musik passieren in der Berufsbildung viele Dinge, deren Einfluss auf den Wirkungsverlauf eingesetzter Maßnahmen unbekannt oder aufgrund der Komplexität schwer abschätzbar ist.

So notwendig der Blick voraus ist, so schwierig ist dieser auf wissenschaftlich möglichst präzise Weise durchzuführen. Generell lässt sich sagen, dass Prognosen vor allem dann vergleichsweise gut sind, wenn sie auf langfristig stabilen Trends aufsetzen und die Einflussfaktoren für diesen Trend zum einen bekannt und zum anderen gut messbar sind. Ein Beispiel hierfür sind demografische Bevölkerungsvorausberechnungen, deren Grundlage drei vergleichsweise stabile Faktoren (Fertilität, Mortalität und Migration) darstellen (vgl. BRETZ 2000). Selbst langfristige Voraussagen sind zuverlässig, soweit es nicht zu dramatischen Katastrophen und damit extremen Veränderungen dieser zentralen Einflussfaktoren kommt (siehe zu den Abweichungen von den Prognosen BRETZ 2002). Für die Berufsbildung lässt sich z. B. die Zahl der Auszubildenden mit einer geringen Streuung für viele Jahre gut voraussagen, und selbst die große Zahl von Zuwanderungen im Jahr 2015 hat an diesen Rahmendaten nur wenig verändert (vgl. BUJARD/DRESCHMITT 2016).

Andere Faktoren wie z. B. die Studierquote der Schulabgänger/-innen und ihre Berufswahl, die Konjunktur der Wirtschaft in einzelnen Branchen (und der damit verbundene Ausbildungsbedarf) sind schon deutlich schwerer voraussagbar, auch wenn es in Deutschland angesichts der guten Datenlage und des fortgeschrittenen Stands der Arbeitsmarkt- und Berufsforschung durchaus seriöse Ansätze gibt (vgl. die umfassende Prognose der Arbeitsmarktentwicklung bis 2030 von VÖGLER-LUDWIG/DÜLL/KRIEHEL 2015). Es sind vor allen Dingen die Funktionalität des Berufsbildungssystems und die anhaltende Wertschätzung der dualen Berufsausbildung in Wirtschaft und Gesellschaft, die hier stabilisierend wirken und eine solche langfristige Vorausschau erst ermöglichen.

Die Berufsbildungszusammenarbeit steht vor größeren Herausforderungen. Die Bedeutung der Rahmenbedingungen für den Erfolg der Berufsbildungszusammenarbeit wird auch von der Bundesregierung in ihrem Strategiepapier zur Berufsbildung betont (vgl. BUNDESREGIERUNG 2013). Das oben angeführte Fallbeispiel der Volksrepublik China mag hierfür als Beleg dienen. In den 35 Jahren der Kooperation mit der HSS hat sich das Land von einem der „Armenhäuser der Welt“ zur führenden Exportnation entwickelt. Die Wachstumsraten waren über viele Jahre stabil zweistellig und die Öffnungspolitik unter Deng Xiaoping führte dazu, dass fast jede international agierende Firma in China Produktionsstätten errichtet hat, Kooperationen mit chinesischen Firmen einging (bzw. eingehen musste) und sowohl für den Weltmarkt als auch zunehmend für den gigantischen chinesischen Binnenmarkt produzierte (vgl. zur Entwicklung Chinas unter Deng Xiaoping VOGEL 2011). Dass z. B. der VW-Konzern die horrenden Strafzahlungen in den USA aufgrund seines Dieselmotors ökonomisch verkraften konnte und trotzdem noch Rekordgewinne erzielte, liegt an der Tatsache, dass VW mittlerweile in China Marktführer ist und der Absatzmarkt dort in ungeheurem Maße boomt (vgl. STOCKMANN/MEYER 2017, S. 112). Mittlerweile gehören einige chinesische Firmen, z. B. der gerade in die Schlagzeilen geratene Mobiltelefon-Hersteller Huawei, zu den Weltmarktführern in ihrem Wirtschaftssegment (vgl. AGTEN/König 2018).

Im Unterschied zu vielen anderen Entwicklungsländern war die Volksrepublik China in diesem Zeitraum politisch stabil und hat in der Berufsbildungspolitik auch keine radikalen Änderungen vorgenommen (vgl. KIRCHBERGER 2014). Dies führt jedoch nicht zwangsläufig zu einer kontinuierlichen Weiterentwicklung der Berufsbildung; es entstehen wirtschaftlich bedingte Wellen, die der Berufsbildungspolitik mal einen höheren und mal einen niedrigeren Stellenwert geben (zur Entwicklung der Berufsbildungspolitik in China siehe STOCKMANN/MEYER 2017, Kap. 2, S. 35ff.).

Insbesondere die fehlende systemische Stabilität stellt sich als Problem für die Berufsbildungskooperation heraus: China hatte zu Anfang kein eigenes Berufsbildungssystem und war an dem Import wesentlicher Komponenten des deutschen dualen Berufsbildungssystems interessiert. Mit dem Boom der Wirtschaft ließ dieses Interesse nach und führte zur Etablierung eines schulischen Systems, in dem Berufsbildung eher als „*Learning by doing*“ in den Betrieben organisiert ist. Die Produktqualität spielte in der Boom-Phase eine untergeordnete Rolle, und die qualitativ hochwertigere Ausbildung an den Kooperationshochschulen der

HSS wurde sozusagen zu einer „Insellösung“ (vgl. STOCKMANN/MEYER 2017, insbesondere S. 198 und S. 224).

Wie kann die Berufsbildungszusammenarbeit in einer solchen Situation wirksam agieren und wissenschaftliche Instrumente zur Wirkungskontrolle einsetzen (vgl. zu dieser Fragestellung ausführlich LI 2017)? Die HSS hat in China über die Zeit ein anpassungsfähiges System des „*muddling through*“ entwickelt, in dem die schnelle Reaktion auf Veränderungen, das offene Ohr für die Partner und die Flexibilität im eigenen Handeln deutlich über einem sorgfältig geplanten und primär an den Wirkungen ausgerichteten Vorgehen gestanden hat. Die Projekte der HSS waren wie kleine Nusschalen im großen chinesischen Meer, die sich der Strömung hingaben und nicht versuchten, im stürmischen Gewässer konsequent Kurs zu halten.

Erfolgreich im Sinne einer Weiterentwicklung der Partnerschaft war diese Strategie sicherlich. Aber wie und in welchem Maße war sie auch wirksam? Hierüber weiß die HSS vergleichsweise wenig, auch wenn sie sich immer wieder um entsprechende Informationen bemüht hat und verschiedene Institute und Personen mit Evaluationen beauftragt hat. Auch das Wirkungsmonitoring hat sich in diesem langen Zeitraum deutlich verbessert, und es konnten Informationen bereitgestellt werden, die zumindest Indizien für einzelne Maßnahmenfolge boten. Ein aus wissenschaftlicher Sicht befriedigendes Monitoring- und Evaluationssystem zur Wirkungsmessung ist jedoch nicht entwickelt worden.

Dies ist aber nicht nur ein primäres Problem der HSS. Die Stiftung ist insgesamt gesehen in der Berufsbildungszusammenarbeit ein kleiner Akteur und andere wie z. B. die Deutsche Gesellschaft für internationale Zusammenarbeit haben, und hatten mit ähnlichen Problemen zu kämpfen. Sie können auch nur bedingt auf vorhandene Konzepte der Wissenschaft zurückgreifen, die für Wirkungsabschätzungen in der Zukunft eingesetzt werden können (vgl. STOCKMANN/MEYER 2018, S. 122f.).

Eine für diesen Zweck entwickelte und in unterschiedlichen Kontexten eingesetzte Methode ist das Delphi-Verfahren (vgl. HÄDER 2014). Allgemein gesagt geht es hier um eine wiederholte Befragung von Fachleuten, bei der dann die unterschiedlichen Meinungen in eine gemeinsame Einschätzung zur zukünftigen Entwicklung überführt werden. Konsensbildung und Prognosefähigkeit als die beiden zentralen Elemente des Delphi-Verfahrens haben dessen Popularität insbesondere in den 1990er-Jahren begründet und zum Einsatz vor allem im Rahmen von Programmentwicklungen geführt. Mittlerweile gibt es viele Varianten dieser Vorgehensweise, die sich teilweise nicht nur auf die Vorausschau, sondern auch auf die Bilanzierung vergangener Entwicklungen beziehen oder die Konsensbildung weniger stark in den Vordergrund rücken (vgl. als ein Überblick zu den verschiedenen Typen und ihren Anwendungen NIEDERBERGER/RENN 2018).

Die Kritik an den Delphi-Verfahren bezieht sich neben diesen Versuchen zur Konsensbildung hauptsächlich auf die Prognosefähigkeit von Expertinnen und Experten. Es lässt sich z. B. generell zeigen, dass Expertinnen und Experten nicht unbedingt weniger Fehler als Laien und Laiinnen machen, sondern eher Fehler anderer Art (vgl. ST. PIERRE/HOFIN-

GER 2014), und dass das Verhältnis zwischen Expertinnen bzw. Experten und Politik durchaus problematisch ist (vgl. BOGNER/TORGERSEN 2005, speziell zur Evaluation siehe EMRICH/GASSMANN/MEYER 2018 und STOCKMANN/MEYER/SCHENKE 2011). Dies gilt insbesondere im Bereich der „prospektiven Politikberatung“, die nur bedingt verlässliche Prognosen liefern kann (vgl. REEZ 2018). Als Fazit aus diesen verschiedenen kritischen Punkten lässt sich sagen, dass Urteile von Fachleuten nur kurzfristig verlässliche Aussagen liefern, Wirkungen neuer Entwicklungen tendenziell überschätzen und oft nicht einheitlich sind.

Wenn es also um die konkrete Voraussage der Zukunft geht, sind Wissenschaftler und Wissenschaftlerinnen am Ende nicht besser als Laien, weil sie sich dazu verleiten lassen, ihr Spezialwissen in einem bestimmten Teilgebiet hervorzuheben und dessen Bedeutung im Gesamtkontext zu überschätzen. Durch Konsensbildung versuchen Delphi-Verfahren dies auszugleichen, was aber nur unter bestimmten Rahmenbedingungen gelingt. Diese und ähnliche Einschränkungen entwerten nicht prinzipiell Expertenurteile und deren Aggregation mittels Delphi-Verfahren, sie sind jedoch in der Anwendung zu beachten.

Aus Sicht der Praxis in der Berufsbildungszusammenarbeit kommt es nicht darauf an, die zukünftige Entwicklung möglichst langfristig vorherzusagen. In den meisten Ländern des globalen Südens muss dies angesichts der hohen Entwicklungsdynamik und der mangelnden Stabilität gesellschaftlicher Institutionen zwangsweise scheitern (vgl. als Beispiel die für den wirtschaftlichen Aufschwung der ostasiatischen „Tigerstaaten“ grundlegenden institutionellen Transformationen mit entsprechend tiefgreifenden Auswirkungen auf die berufliche Bildung, EBNER 2014). Der hohe Aufwand des Delphi-Verfahrens zur Gewinnung verlässlicher Prognosen scheint angesichts dieser Umstände kaum gerechtfertigt.

Das Centrum für Evaluation erprobt deshalb seit einigen Jahren in Costa Rica eine vereinfachte Variante des Delphi-Verfahrens, das Expertenpanel. Diese Vorgehensweise unterscheidet sich in mehrerlei Hinsicht von den beschriebenen methodischen Ansätzen:

- ▶ Die Expertinnen und Experten werden jeweils zu Veränderungen in den vergangenen Jahren und erwarteten Entwicklungen im nächsten Jahr befragt, d. h., es geht nicht um die Generierung langfristiger Prognosen, sondern um kurzfristige und überschaubare Zeiträume.
- ▶ Es werden immer wieder die gleichen Expertinnen und Experten befragt, welche dann auch Gründe für Abweichungen von ihren Erwartungen nennen können. Hierdurch werden nicht nur Trends, sondern auch (potenzielle) Einflussfaktoren erhoben.
- ▶ Allgemeine Prognosen und retrospektive Betrachtungen werden in einen Zusammenhang mit konkreten Projektmaßnahmen und Interventionen gebracht. Dies geschieht einerseits durch die Erhebung von Einflussfaktoren bei den Expertinnen und Experten, die dann zu Aktivitäten des Projektes in Bezug gesetzt werden können. Andererseits können die Expertinnen und Experten selbst nach einer Einschätzung der Effekte bestimmter Maßnahmen gefragt werden.

- ▶ Das Expertenpanel wird zur Hypothesengenerierung und nicht zur Wirkungsprüfung eingesetzt, d. h., die Aussagen der Expertinnen und Experten werden nicht als Abbildung der Realität, sondern als Annahmen verstanden, die über weitere Datenquellen und mit anderen Methoden gefestigt oder widerlegt werden können. Hinsichtlich des prognostischen Teils werden sie als eine – nicht aber die einzige – Grundlage für (kurzfristig ausgerichtete) Steuerungsentscheidungen verwendet. Hierdurch soll vor allem verhindert werden, dass bestimmte Entwicklungstrends von der Programmleitung übersehen oder falsch eingeschätzt werden.

Das Expertenpanel hat sich als praktikables Verfahren über viele Jahre bewährt, und es konnten eine Reihe methodischer Lehren aus seiner Anwendung im Kontext der Berufsbildungszusammenarbeit gewonnen werden. Eine Publikation dieser Erfahrungen ist derzeit in Vorbereitung.

Das Expertenpanel ist sicherlich nicht die Lösung für alle beschriebenen methodischen und praktischen Probleme der Berufsbildungszusammenarbeit hinsichtlich der Erfassung von Wirkungen und ihrer Verläufe. Es ist aber im Kontext anwendungsorientierter methodischer Innovationen zur Unterstützung eines langfristig ausgerichteten wirkungsorientierten Managements komplexer Vorhaben zu sehen, welche die Limitierung eines primär wirkungstestenden experimentellen Denkens überwinden wollen (weitere Beispiele finden sich im Komplexitätsmanagement von Unternehmen, vgl. SCHOENEBERG 2014, oder der Nexus-Initiative zur Steuerung von nachhaltiger Entwicklung, vgl. STEIN/BARRON/MOSS 2014). Hinzu kommen die neuen technischen Möglichkeiten zur Erschließung des Internets als Datenquelle für Monitoring und Evaluation, die unter der Rubrik „*Big Data*“ gegenwärtig diskutiert und eingesetzt werden (vgl. PETERSEN/BREUL 2017). In diesem Kontext werden sicher auch experimentelle Designs in Zukunft ihren Platz finden, allerdings eher ergänzend und nicht mit einem Führungsanspruch, wie ihn der Begriff „Goldstandard“ impliziert.

5 Zusammenfassung und Ausblick

Dieser Beitrag hat aus einer Praxisperspektive der Berufsbildungszusammenarbeit die experimentelle Methode als Mittel zur Wirkungsmessung infrage gestellt und sich dabei der Analogie zur harmonischen Zusammenarbeit einzelner Instrumente in einer Sinfonie sowie den dort auftretenden Schwierigkeiten einer umfassenden Wirkungsbeurteilung bedient. Im Wesentlichen sind drei Argumente angeführt worden:

1. Es geht in der Berufsbildungszusammenarbeit selten um die Frage, ob bestimmte Maßnahmen oder ganze Programme überhaupt wirken, sondern eher um die Wirkungsverläufe in unterschiedlichen Kontexten. Es sind Fragen wie: Welche Faktoren beschleunigen oder verlangsamen den Eintritt von Wirkungen? Wie lange bleiben die Wirkungen auf einem bestimmten Niveau und wie weit sinken die Effekte ab? Was passiert, wenn Projektinterventionen beendet werden – bleiben die Maßnahmenwirkungen dann auf

demselben Niveau bestehen oder sinken sie ab (und wenn ja, bis auf welches Niveau)? Da es sich zumeist um bereits erprobte Instrumente und Maßnahmen handelt, ist die grundsätzliche Frage, ob sie überhaupt wirken, eher von untergeordnetem Interesse.

2. Experimente sind nicht in der Lage, die Komplexität der Berufsbildungszusammenarbeit insgesamt zu erfassen. Als reduktionistische Verfahren sind sie eher geeignet, eng begrenzte und isolierte Elemente zu testen und nicht umfassende Programme mit einer Vielzahl unterschiedlicher Wirkungszusammenhänge. Zur Erfassung der Wirkungen komplexer Systeme braucht es dagegen eine holistische Betrachtungsweise, welche die Möglichkeiten eines experimentellen Designs bei Weitem übersteigt.
3. Schließlich ist besonders im Kontext von Nichtregierungsorganisationen die Zusammenarbeit nicht eindeutig zeitlich begrenzt und kann deshalb schwer mittels einer Vorher-nachher-Messung erfasst werden. Wirkungsverläufe in zeitlich unbefristeten Vorhaben sind nur durch kontinuierliche Messungen abzubilden. Insbesondere die Terminierung der Messzeitpunkte für die Durchführung eines Experiments stellt eine Herausforderung dar, welche aufgrund fehlender theoretischer Grundlagen in vielen Fällen kaum adäquat bewältigt werden kann.

Experimente sind deshalb mitnichten der „Goldstandard“ zur Wirkungsmessung in der Berufsbildungszusammenarbeit, die Praxis braucht im Gegenteil andere und ihren Erfordernissen angemessenere Vorgehensweisen. Entsprechend den hier beschriebenen Anforderungen geht es um Verfahren, die a) in der Lage sind, Zukunftsabschätzungen zu Wirkungspotenzialen vorzunehmen, b) Wirkungshypothesen über eingetretene Effekte in der Vergangenheit generieren können und c) kontinuierlich vergleichbare Messdaten produzieren, welche die Wirkungsverläufe abdecken und dadurch steuerbar machen. All dies ist mit Experimenten nicht leistbar, weshalb die Berufsbildungszusammenarbeit von der Berufsbildungsforschung neue und ihren Bedürfnissen besser angepasste Methoden benötigt. Hieran sollte systematischer als in der Vergangenheit gearbeitet werden.

Literatur

- AGTEN, Sven; König, Thomas: Die echte Innovation kommt aus dem Osten – Made in China 2.0. In: AGTEN, Sven; KÖNIG, Thomas (Hrsg.): So schafft man China. Wie Sie Business und Alltag meistern. Wiesbaden 2018, S. 199–226
- ARNOLD, Rolf; FEDER, Fritz: Berufsbildung in der Entwicklungszusammenarbeit. In: Bildung und Erziehung 58 (2005) 4, S. 433–447
- BOGNER, Alexander; TORGERSEN, Helge (Hrsg.): Wozu Experten? Ambivalenzen der Beziehung von Wissenschaft und Politik. Wiesbaden 2005
- BOSCH, Gerhard: Facharbeit, Berufe und berufliche Arbeitsmärkte. In: WSI-Mitteilungen (2014) 1, S. 5–13

- BRETZ, Manfred: Methoden der Bevölkerungsvorausberechnung. In: MUELLER, Ulrich; NAUCK, Bernhard; DIEKMANN, Andreas (Hrsg.): Handbuch der Demographie. Bd. 1. Berlin, Heidelberg 2000, S. 643–681
- BRETZ, Manfred: Treffsicherheit von Bevölkerungsvorausberechnungen. In: BUNDESINSTITUT FÜR BEVÖLKERUNGSFORSCHUNG (Hrsg.): Demographische Vorausschätzungen Grenzen und Möglichkeiten, Methoden und Ziele: Vorträge auf der gemeinsamen Sitzung des Arbeitskreises „Bevölkerungswissenschaftliche Methoden“ der Deutschen Gesellschaft für Bevölkerungswissenschaft und des Ausschusses für Regionalstatistik der Deutschen Statistischen Gesellschaft im Rahmen der Statistischen Woche in Nürnberg am 26./27. September 2000. Materialien zur Bevölkerungswissenschaft 104. Wiesbaden 2002 – URL: <https://nbn-resolving.org/urn:nbn:de:0168-ssaoar-334028> (Zugriff: 02.12.2019)
- BUJARD, Martin; DRESCHMITT, Kai: Szenarien der Bevölkerungsentwicklung bis 2060. Wie beeinflussen Migration und Geburten Deutschlands Zukunft? In: Gesellschaft. Wirtschaft. Politik (GWP) (2016) 3, S. 333–346 – URL: <https://doi.org/10.3224/gwp.v65i3.07> (Zugriff: 02.12.2019)
- BUNDESREGIERUNG (Hrsg.): Strategiepapier der Bundesregierung zur internationalen Berufsbildungszusammenarbeit aus einer Hand. Drucksache 17/14352 des Deutschen Bundestags vom 05.07.2013
- DENZIN, Norman K.: The Experiment and its Variations. In: DENZIN, Norman K. (Hrsg.): The Research Act. New York 2009, S. 147–163
- EBNER, Alexander: Die industriepolitische Transformation der ostasiatischen Entwicklungsstaaten. In: Nölke, Andreas; MAY, Christian; CLAAR, Simone (Hrsg.): Die großen Schwellenländer. Ursachen und Folgen ihres Aufstiegs für die Weltwirtschaft. Wiesbaden 2014, S. 135–151
- EMRICH, Eike; GASSMANN, Freya; MEYER, Wolfgang: Geliefert wie bestellt: Wa(h)re Wissenschaft? In: POFERL, Angelika; KELLER, Reiner (Hrsg.): Wissenskulturen der Soziologie. Weinheim, Basel 2018, S. 113–137
- GESSLER, Michael: Educational Transfer as Transformation: A Case Study about the Emergence and Implementation of Dual Apprenticeship Structures in a German Automotive Transplant in the United States. In: Vocations and Learning (2017) 10, S. 71–99
- GESSLER, Michael; KÜHN, Kristina: Geschäftsmodellentwicklung und/oder Internationale Berufsbildungszusammenarbeit? In: DLR PROJEKTRÄGER (Hrsg.): Berufsbildung International Geschäftsmodellentwicklung. Bonn 2019, S. 35–36
- Häder, Michael: Delphi-Befragungen. Ein Arbeitsbuch. 3. Aufl. Wiesbaden 2014
- HEITMANN, Werner: 60 Jahre internationale Förderung der Berufsbildung im Überblick: Veränderungsprozesse in der Wertschöpfungskette für Berufsbildung in der staatlichen Berufsbildungszusammenarbeit im Ausland. In: GESSLER, Michael; FUCHS, Martina; PILZ, Matthias (Hrsg.): Konzepte und Wirkungen des Transfers Dualer Berufsausbildung. Wiesbaden 2019, S. 59–118

- KAPTCHUK, Ted J.: The double-blind, randomized, placebo-controlled trial: Gold standard or golden calf? In: *Journal of Clinical Epidemiology* 54 (2001) 6, S. 541–549
- KIRCHBERGER, Sarah: Lernfähiger Leninismus? Das politische System der Volksrepublik China. In: FISCHER, Doris; MÜLLER-HOFSTEDT, Christoph (Hrsg.): *Länderbericht China*. Bonn 2014, S. 251–288
- LANGTHALER, Margarita: Die entwicklungspolitischen Wirkungen von Bildung. Rezentere Forschungsergebnisse und ihre Implikationen für die Entwicklungszusammenarbeit. In: *Policy Note 04/2013* (hrsg. v. Österreichische Forschungsstiftung für Internationale Entwicklung) – URL: https://www.oefse.at/fileadmin/content/Downloads/Publikationen/Policynote/PN4_Wirkungen_von_Bildung.pdf (Zugriff: 02.12.2019)
- LENZEN, Marcus: Die Rolle der NGOs in der Entwicklungszusammenarbeit. Münsteraner Diskussionspapiere zum Nonprofit-Sektor 17. Münster 2001 – URL: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-361881> (Zugriff: 02.12.2019)
- LI, Junmin: *Policy-Transfer von deutschen Evaluationskonzepten der Berufsbildung nach China. Eine Analyse am Beispiel des Peer-Review-Verfahrens*. Wiesbaden 2017
- MEYER, Wolfgang: Einführung in die Grundlagen der Entwicklung von Indikatoren. In: WROBLEWSKI, Angela; KELLE, Udo; REITH, Florian (Hrsg.): *Gleichstellung messbar machen. Grundlagen und Anwendungen von Gender- und Gleichstellungsindikatoren*. Wiesbaden 2016, S. 15–38
- NIEDERBERGER, Marlen; RENN, Ortwin: *Das Gruppendelphi-Verfahren. Vom Konzept bis zur Anwendung*. Wiesbaden 2018
- PETERSSON, Gustav; BREUL, Jonathan D. (Hrsg.): *Cyber Society, Big Data, and Evaluation*. New York 2017
- PINCH, Trevor J.; TROCCO, Frank: *Analog days: The invention and impact of the Moog synthesizer*. Cambridge 2009
- PRESCOTT-CLEMENTS, Linda u. a.: Evidence for validity within workplace assessment: the Longitudinal Evaluation of Performance (LEP). In: *Medical Education* 42 (2008) 5, S. 488–495
- REEZ, Norbert: Prospektive Politikberatung. In: *SIRIUS – Zeitschrift für strategische Analysen* (2018) 2/3, S. 285–287 – URL: <https://doi.org/10.1515/sirius-2018-3009> (Zugriff: 02.12.2019)
- SCHELTEN, Andreas: Berufsmotorisches Lernen in der Berufsbildung. In: BONZ, Bernhard (Hrsg.): *Didaktik und Methodik der Berufsbildung*. Baltmannsweiler 2009, S. 135–151
- SCHIFFNER, Wolfgang: *Lexikon Tontechnik*, Kassel 2003
- SCHOENEBERG, Klaus-Peter (Hrsg.): *Komplexitätsmanagement in Unternehmen. Herausforderungen im Umgang mit Dynamik, Unsicherheit und Komplexität meistern*. Wiesbaden 2014
- STEIN, Christian; BARRON, Jennie; MOSS, Timothy: *Governance of the nexus: from buzz words to a strategic action perspective*. Nexus Network Think Piece Series, Paper 003. Brighton 2014

- STOCKMANN, Reinhard u. a.: Wirksamkeit Deutscher Berufsbildungszusammenarbeit. Ein Vergleich staatlicher und nicht-staatlicher Programme in der Volksrepublik China. Op-laden 2000
- STOCKMANN, Reinhard; MEYER, Wolfgang; SCHENKE, Heike: Unabhängigkeit von Evaluationen. In: Zeitschrift für Evaluation 10 (2011) 1, S. 39–67
- STOCKMANN, Reinhard; MEYER, Wolfgang: Chinas Berufsbildung im Wandel. 30 Jahre Entwicklungszusammenarbeit mit der Hanns-Seidel-Stiftung. Münster 2017
- STOCKMANN, Reinhard; MEYER, Wolfgang: Berufsbildungsforschung in der Entwicklungszusammenarbeit. In: RAUNER, Felix; GROLLMANN, Philipp (Hrsg.): Handbuch Berufsbildungsforschung, 3. Aufl. Bielefeld 2018, S. 117–126
- ST. PIERRE, Michael; HOFINGER, Gesine: Fehler und Fehlerursachen. In: ST. PIERRE, Michael; HOFINGER, Gesine (Hrsg.): Human Factors und Patientensicherheit in der Akutmedizin. 3. Aufl. Berlin, Heidelberg 2014, S. 45–68.
- VOGEL, Ezra F.: Deng Xiaoping and the Transformation of China. Cambridge, London 2011 – URL: http://gelora45.com/news/DengXiaoping_TransformationOfChina_Vogel.pdf (Zugriff: 02.12.2019)
- VOGLER-LUDWIG, Kurt; DÜLL, Nicola; KRIEHEL, Ben: Arbeitsmarkt 2030. Die Bedeutung der Zuwanderung für Beschäftigung und Wachstum. Prognose 2014. Bielefeld 2015 – URL: <http://www.oapen.org/search?identifier=640934> (Zugriff: 02.12.2019)
- WINDELBAND, Lars: Zukunft der Facharbeit im Zeitalter „Industrie 4.0“. In: Journal of Technical Education 2 (2014) 2, S. 138–160
- WOLTER, Marc Ingo u. a.: Industrie 4.0 und die Folgen für Arbeitsmarkt und Wirtschaft: Szenario-Rechnungen im Rahmen der BIBB-IAB-Qualifikations- und Berufsfeldprojektionen. IAB-Forschungsbericht 8/2015. Nürnberg 2015

Lydia Schulze Heuling, Steffen Wild, Ernst Deuer

► **Längere Schulzeit – andere Ansprüche? Wie die Beschulungsdauer die Prioritäten bei der Arbeitgeberwahl beeinflusst**

Die vorliegende Studie untersucht, ob sich die Beschulungsdauer (G8/G9) in den Ansprüchen niederschlägt, die Schülerinnen und Schüler später als Studierende an potenzielle Arbeitgeber stellen. Befragt wurden 299 dual Studierende der Dualen Hochschule Baden-Württemberg, die für Unternehmen begehrte *high potentials* sind. Die Ergebnisse des *Propensity Score Matchings* zeigen, dass Abiturientinnen und Abiturienten des achtjährigen Gymnasiums (G8) vor allem Wert auf Sicherheit und Karriereperspektiven legen. Im Gegensatz dazu ist Studierenden, die das neunjährige Gymnasium (G9) absolviert haben, die Förderung ihres Innovationspotenzials wichtiger. Die durchgeführten Sensitivitätsanalysen weisen auf weitestgehend robuste Ergebnisse.

1 Hintergründe und Forschungsstand

Nach einer kurzen Darlegung der Schulzeitdebatte bis zum Abitur sowie der Arbeitgeberattraktivität und des dualen Studiums analysieren wir, wie die Beschulungsdauer die studentische Meinung über einen attraktiven Arbeitsplatz beeinflusst.

1.1 Schulzeitdebatte

Die länderabhängige Einführung einer verkürzten Schulzeit bis zum Abitur von zwölf Jahren (G8) hat eine kontroverse Debatte ausgelöst (vgl. HUEBENER/MARCUS 2015a). Einige Bundesländer, darunter Niedersachsen und Bayern, haben jüngst die eingeführte Reform wieder rückgängig gemacht. Dagegen hat dieses Schulmodell in den neuen Bundesländern Tradition und erfährt großen Rückhalt (vgl. WÖSSMANN u. a. 2014).

Im argumentativen Fokus für eine Schulzeitverkürzung liegen vor allem die Senkung des Erwerbseintrittsalters und die Beschleunigung der Arbeitsmarktfähigkeit (*employability*) von Abiturientinnen und Abiturienten (vgl. HUEBENER/MARCUS 2015b; KLEMM 2008). In diesem Zusammenhang stellen sich dann allerdings auch Fragen nach mittel- und lang-

fristigen Folgen. Dieser Beitrag untersucht, ob die Beschleunigung und inhaltliche Verdichtung des Oberstufencurriculums durch eine Schulzeitverkürzung um ein Jahr Auswirkungen auf die Ansprüche hat, die Absolventinnen und Absolventen an den zukünftigen Arbeitsplatz stellen. Falls dies der Fall ist, bestehen strukturelle und inhaltliche Herausforderungen sowohl für Schulen als auch für Betriebe, die sich um Auszubildende oder qualifizierte Studienabsolventinnen und -absolventen bemühen, sowie schließlich auch für Hochschulen mit berufsqualifizierenden Studiengängen.

Ein historischer Blick auf die Entwicklung der Regelschulzeit bis zum Abitur macht deutlich, dass diese Angelegenheit vielfach erörtert und angepasst wurde (vgl. KÜHN u. a. 2013). Die deutsch-deutsche Wiedervereinigung machte in jüngerer Zeit eine erneute Auseinandersetzung mit der Schulzeit bis zum Abitur notwendig, denn in den Ländern der ehemaligen DDR war das Gymnasium achtjährig (G8), in den Ländern der BRD neunjährig (G9). Wer sich wem anzupassen hatte, schien dabei keine Frage: Den ostdeutschen Ländern wurde im Rahmen der Übergangvereinbarungen eine Übergangszeit zur Umstellung von zwölf auf 13 Schuljahre eingeräumt (vgl. KÜHN u. a. 2013). Gleichzeitig jedoch setzte ein sich neuformierender, internationalisierter Bildungsfokus eine gegenläufige Bewegung in Gang, woraufhin einige westdeutsche Bundesländer die Schulzeit in der Sekundarstufe II um ein Jahr, auf zwölf Schuljahre, verkürzten. Das Saarland als erstes westdeutsches Bundesland führte das G8 mit dem Schuljahr 2001/2002 ein (vgl. KÜHN 2015). Nordrhein-Westfalen folgte 2004. Das Saarland begründete die Verkürzung vor allem mit der geografischen Nähe zu Frankreich und einer daraus resultierenden Benachteiligung saarländischer Absolventinnen und Absolventen gegenüber französischen, die ein Jahr früher ihren Abschluss erhalten. Nordrhein-Westfalen sprach allgemeiner von einem verantwortungsvolleren Umgang mit der Lebenszeit unserer Kinder (vgl. MINISTERIUM FÜR SCHULE UND WEITERBILDUNG DES LANDES NORDRHEIN-WESTFALEN 2016).

Huebener und Marcus (2015b) identifizieren drei zentrale Argumente für eine Schulzeitverkürzung von G9 auf G8:

1. Senkung des im internationalen Vergleich hohen Eintrittsalters der Jugendlichen in den Arbeitsmarkt;
2. Steigerung der internationalen Wettbewerbsfähigkeit deutscher Abiturientinnen und Abiturienten und
3. eine Entlastung der sozialen Sicherungssysteme bzw. des Arbeitsmarkts aufgrund des Fachkräftemangels im Hinblick auf die alternde Gesellschaft.

Schavan und Ahnen (2001) erklären darüber hinaus, dass die Reform einem vermeintlich schnelleren Reifeprozess heutiger Jugendlicher entgegenkommt. Weiterhin werden in der Literatur eine unnötige Stagnation durch zu lange Orientierungsphasen, umfangreiche unterrichtliche Wiederholungsphasen, ausgedehnte Abiturvorbereitung und -prüfungszeiten sowie ohnehin hohe Unterrichtsausfälle für eine Schulzeitverkürzung genannt. Dagegen argumentieren die G8-Kritikerinnen und -Kritiker, dass durch die Verkürzung der gymnasialen

Schulzeit ein eher oberflächliches Lernen gefördert werde und die Persönlichkeitsentwicklung, die Entwicklung moralischen Bewusstseins sowie die psychosoziale Entwicklung Zeit brauchen (vgl. KÜHN u. a. 2013). Kühn u. a. (2013) fassen die Argumente für G9 folgendermaßen zusammen: Die zeitliche Verkürzung um ein Schuljahr nimmt wichtige Zeit für die individuelle berufliche Orientierung, für Auslandsschuljahre, Klassenfahrten und Besuche außerschulischer Lernorte. Schülerinnen und Schüler stehen laut Kühn u. a. (2013) im achtjährigen Gymnasium zudem unnötig unter erhöhtem Leistungsdruck und Stress und die Durchlässigkeit des Bildungssystems für alternative Bildungsbiografien nimmt ab.

Bei Schavan/Ahnen (2001) heißt es hingegen, dass eine Schulzeitverkürzung gerade wieder eine Besinnung und Konzentration auf grundlegende Bildungsinhalte, die Stabilisierung fachlicher Grundlagen und methodischer Kompetenzen sowie die Entwicklung von Orientierung und Urteilsvermögen möglich macht.

Seit der Finanzkrise im Jahr 2009 führen Befürworterinnen und Befürworter der kürzeren Schulzeit bis zum Abitur verstärkt auch ökonomische bzw. finanzpolitische Gründe an. So könnten durch die Streichung der Unterrichtszeit eines ganzen Schuljahres zahlreiche Stellen für Lehrkräfte eingespart werden. Dagegen gehalten wird vor allem, dass so eingesparte Ressourcen an anderer Stelle gebraucht würden, etwa für die Überarbeitung von Schulbüchern und Unterrichtsmaterialien, und dass Einsparungen auf diesem Wege nicht stichhaltig sind (vgl. KÜHN u. a. 2013).

Aufgrund verschiedener Entwicklungen wird die gegenwärtige Situation in den 16 Bundesländern als Flickenteppich bezeichnet (vgl. KULTUSMINISTERKONFERENZ 2016): In den neuen Bundesländern besteht ausnahmslos das G8, und es gibt gegenwärtig keine Anstrengung, zu G9 überzugehen. Niedersachsen kehrte mit Beginn des Schuljahres 2015/2016 zu G9 zurück. Am 5. April 2017 erklärte auch Bayern die Rückkehr zum G9 ab dem Schuljahr 2018/2019 (vgl. BAYERISCHES STAATSMINISTERIUM FÜR BILDUNG UND KULTUS, WISSENSCHAFT UND KUNST 2017). Rheinland-Pfalz führte G8 nie richtig ein. Die verbleibenden Bundesländer befinden sich in einer Art Schwebezustand, der dadurch gekennzeichnet ist, dass der G8-Zug zwar offiziell eingeführt ist, allerdings das G9 unter besonderen Umständen beibehalten werden kann (vgl. AUTORENGRUPPE BILDUNGSBERICHTERSTATTUNG 2016; BURCHARD 2016).

Die empirischen Befunde zur Schulzeitdebatte deuten darauf hin, dass die Reform ambivalent zu bewerten ist (vgl. WÖSSMANN u. a. 2014). Es gibt Hinweise, dass sich das Stress-Argument der G8-Gegnerinnen und G8-Gegner nicht bestätigt. Minkley u. a. (2015) fanden heraus, dass sich Schülerinnen und Schüler in G8-Gymnasien nicht gestresster fühlen als die Vergleichsgruppe des G9-Zugs. Milde-Busch u. a. (2010) hingegen diagnostizieren, dass die den Schülerinnen und Schülern zur Verfügung stehende Freizeit nicht zur Regeneration und ausgleichenden Freizeitgestaltung ausreicht. Studienergebnisse zum Einfluss der Schulzeit auf Persönlichkeitseigenschaften wie die *Big Five* (vgl. OSTENDORF 1990) oder die Selbstbeherrschung (vgl. MISCHEL 2014) sind ebenfalls heterogen und kommen zu unterschiedlichen Schlüssen. Thiel, Thomsen und Büttner (2014) stellen keine Effekte der Schulzeitverkürzung auf die *Big Five*, die Selbstkontrolle und Kontrollüberzeugung fest. Im Gegensatz dazu

kommen Dahmann und Anger (2014) zu dem Schluss, dass G8-Absolventinnen und -Absolventen extrovertierter, aber emotional weniger stabil sind als G9-Absolventinnen und -Absolventen. Huebener und Marcus (2015b) schließlich zeigen, dass sich nach Einführung des G8 das Durchschnittsalter der Abiturientinnen und Abiturienten um nur etwa zehn Monate reduziert, was sie vor allem auf die steigende Zahl von Klassenwiederholungen zurückführen. Ein Vergleich der Schulleistungen zwischen G8 und G9 weist ebenfalls in keine eindeutige Richtung (vgl. HUEBENER/MARCUS 2015a). Die aktuelle Studie von Büttner und Thomsen (2015) ermittelte, basierend auf der Untersuchungspopulation des Doppeljahrgangs in Sachsen-Anhalt, dass sich G8 negativ auf die Mathematiknote auswirkt. Weitere Studien der Kultusministerien bzw. statistischen Ämter ergeben insgesamt, dass Leistungsänderungen in unterschiedliche Richtungen weisen und widersprüchlich sind.

1.2 Auswirkungen von G8 und G9 auf das Studierverhalten

Wenden wir unseren Blick auf die Zeit nach dem Erwerb der Hochschulreife – auf das Studium. Bisherige Untersuchungen zeigen, dass sich G8- und G9-Studierende in ihrer Studierfähigkeit, der Motivation im Studium und der Studienabbruchwahrscheinlichkeit nicht unterscheiden (vgl. MEYER/THOMSEN 2014). Ähnlich kommen Derr, Hübl und Zaki (2013) beim Vergleich von G8- und G9-Abiturientinnen und -Abiturienten im Rahmen eines Matheignungstests an der Dualen Hochschule Baden-Württemberg (DHBW) zu dem Ergebnis, dass es keine Unterscheide gibt. Dagegen finden Pustelnik und Halverscheid (2013) etwas höhere Leistungen in Mathematik Klausuren an der Universität Göttingen zugunsten der G8-Absolventinnen und G8-Absolventen. Die Forschungsbefunde zu den Folgen der Schulungszeiten G8 und G9 für die Hochschulen sind bisher nur wenig erforscht und geben, ähnlich zu den oben skizzierten Schulstudien, kein eindeutiges Bild. Während Ladenthin, Krautz und Nostadt (2016) eine hohe Studienabbruchquote trotz guter Abiturnoten, einen Mangel kognitiver Entwicklung zur angemessenen Aufnahme komplexer Prozesse sowie eine Unfähigkeit zum wissenschaftlichen Diskurs und mangelnde Lösungskompetenz in moralischen Konfliktsituationen als Folge der verkürzten Regelschulzeit zu identifizieren wissen, weist Kühn (2014) mit ihrer umfangreichen Studie nach, dass keine faktischen Unterschiede zwischen den Absolventinnen und Absolventen der beiden Schulformen bestehen. Insgesamt muss jedoch festgehalten werden, dass der Forschungsstand zur Auswirkung der Regelschulzeit bis zum Abitur wenig systematisch und umfangreich ist und die Befundlage heterogen. Insbesondere ist wenig bekannt über die kleine Gruppe der dual Studierenden bzw. Absolventinnen und Absolventen dualer Studiengänge. Diese Gruppe nimmt die vorliegende Studie in den Blick und untersucht, ob sich der Besuch eines acht- oder neunjährigen Gymnasiums auf die Einschätzung der Arbeitgeberattraktivität auswirkt.

Kriterien zu untersuchen, die G9- und G8-Absolventen/Absolventinnen an zukünftige Arbeitgeber stellen, ist insbesondere virulent, weil die Erhöhung der Wettbewerbsfähigkeit von Abiturientinnen und Abiturienten aus Deutschland eine zentrale Begründung der Schulzeitverkürzung war. Um jedoch qualifizierte Mitarbeiterinnen und Mitarbeiter, gerade auch

zukünftige *high potentials*, für sich gewinnen zu können und beispielsweise nicht an konkurrierende Firmen zu verlieren, müssen Arbeitgeber für Absolvierende attraktiv sein (vgl. PARMENT 2013). Dies gilt insbesondere für Studierende dualer Studiengänge. Ein duales Studium bedeutet, zeitgleich in einer koordinierten betrieblichen und einer hochschulischen Ausbildung zu sein. Diese Doppelqualifikation macht die Absolventinnen und Absolventen solcher Ausbildungsformate in mehrfacher Hinsicht besonders interessant für Betriebe (vgl. DEUER/TRÄGER 2015). Aufschluss über Kriterien zu erhalten, die dual Studierende an zukünftige Arbeitgeber anlegen, ist jedoch nicht nur für Betriebe wichtig. Auch Studierenden bietet eine Reflexion der eigenen Ansprüche gegenüber zukünftigen Arbeitgebern eine wichtige Orientierungsgrundlage. Ferner besitzen dual Studierende bereits erste Berufserfahrungen aufgrund der berufspraktischen Zeit im Studium, was die Aussagekraft dieser Studie erhöht.

1.3 Arbeitgeberattraktivität

Betriebswirtschaftliche und wirtschaftspsychologische Studien weisen immer wieder darauf hin, dass Betriebe, die junge Mitarbeiterinnen und Mitarbeiter gewinnen wollen, Kenntnisse über die Wünsche, Erwartungen und Bedürfnisse der jüngeren Generationen benötigen. Unter Umständen kommen die Unternehmen im *war of talents* (vgl. HOLSTE 2012) nicht daran vorbei, die Arbeitsbedingungen und Angebote für Arbeitnehmende an deren Erwartungen anzupassen. Einige Betriebe bzw. Konzerne gehen sogar so weit, ein sogenanntes *Employer Branding* (Arbeitgebermarkenbildung) aufzubauen, um junge Talente für sich zu gewinnen und an sich zu binden (vgl. KRAUSE 2015; RUTHUS 2013).

Eine theoretische Fundierung der Arbeitgeberattraktivität wird aus unterschiedlichen Richtungen vorgenommen. Manchmal wird auf Modelle der Motivationsforschung aus Arbeits- und Organisationspsychologie zurückgegriffen. Ruthus (2013) beispielsweise greift in seiner Konzeptionierung von Selbstverwirklichung auf die ERG-Theorie (*Existence needs – Relatedness needs – Growth needs*) von Alderfer (1972) zurück. Solche Theorieansätze greifen jedoch nur sehr begrenzt spezifische Fragestellungen zur Arbeitgeberattraktivität auf und erklären diese nur eingeschränkt. Motivationsforschung hingegen fokussiert eher auf Aspekte der Effizienz von Arbeitnehmenden als auf motivationale Entscheidungsgrundlagen bei der Wahl einer bestimmten Arbeitgeberin bzw. eines bestimmten Arbeitgebers. Andere Studien konfrontieren Studierende mit kaum differenzierten Fragen, die oftmals auch nur unzureichend theoretische Konstrukte referenzieren. So etwa Sommer (2016), der Studierende an Hochschulen befragte, welches Unternehmen die attraktivsten Arbeitsplätze habe. Die Antworten auf so pauschale Fragen enthalten ein nur sehr eingeschränktes Aufklärungspotenzial, zumal die befragte Population noch keine Berufserfahrung vorweisen kann.

Weit verbreitet ist der Ansatz, die Beurteilung von Unternehmen aus der Sicht (zukünftiger) Beschäftigter über die generationenspezifischen Werte und Einstellungen zu erforschen (vgl. PARMENT 2013; KRAUSE 2015; LEBRENTZ/REGNET 2013). Dazu wird auf charakteristische

Merkmale der Bevölkerungskohorten zurückgegriffen: die Baby Boomer als stark leistungsorientiert und mit ausgeprägtem Pflichtbewusstsein, die Generation X als nach Selbstständigkeit strebend und ehrgeizig oder die Generation Y als medienaffin und gut ausgebildet (vgl. RUTHUS 2013; DEUTSCHE GESELLSCHAFT FÜR PERSONALFÜHRUNG 2011; HOLSTE 2012). Aktuelle Studien zufolge legt die jüngere und jüngste Generation der Hochschulabsolvierenden Wert auf a) Spaß an der Arbeit und herausfordernde Aufgaben, b) Weiterbildungsmöglichkeiten, c) die Wertigkeit der im Betrieb hergestellten oder verkauften Produkte bzw. Dienstleistungen und auf d) Sinnerfüllung (vgl. RUTHUS 2013; DEUTSCHE GESELLSCHAFT FÜR PERSONALFÜHRUNG 2011). Die Vereinbarkeit von Familie und Beruf (vgl. RAMM u. a. 2014) sowie flexible Arbeitszeiten oder Arbeitsphasen, die sich dem persönlichen Lebensrhythmus anpassen lassen (vgl. LEVEN/QUENZEL/HURRELMANN 2015), sind darüber hinaus zentrale Ansprüche, die Studierende und junge Hochschulabsolventinnen und Hochschulabsolventen an Unternehmen stellen. Laut der jüngsten Shell-Jugendstudie ist das Sicherheitspotenzial eines Arbeitsplatzes, gefolgt von Sinnerfüllung und gesellschaftlicher Relevanz (vgl. SHELL DEUTSCHLAND HOLDING 2015), das wichtigste Kriterium junger Menschen an den Arbeitsplatz. Karrierechancen sind laut dieser Studie eher nachgeordnet (vgl. LEVEN u. a. 2015). Andere Untersuchungen ergänzen die Befundlage und zeigen, dass Forschungsaktivitäten und wissenschaftliche Tätigkeiten ebenfalls nachgeordnete Kriterien von jüngeren Absolventinnen und Absolventen sowie Berufsanfängerinnen und Berufsanfängern für die Betriebswahl sind (vgl. RAMM u. a. 2014). Die Bewertung dieser Kriterien ist in den letzten Jahren sogar noch leicht gesunken (vgl. MULTURS/BARGEL/RAMM 2008).

1.4 Dual Studierende

Das duale Studium findet in den letzten Jahren zunehmenden Zuspruch und weist steigende Studierendenzahlen auf. Waren es für das Jahr 2004 noch rund 41.000 Studierende, stieg die Zahl auf fast 60.000 Studierende im Jahr 2011 an (vgl. AUSBILDUNGPLUS 2014). Eine Schätzung geht im Jahr 2019 von rund 108.000 Studierenden aus (vgl. HOFMANN u. a. 2019). Etwa ein Drittel dieser Studierenden war an der DHBW (34.390 Studierende) eingeschrieben (vgl. STATISTISCHES LANDESAMT BADEN-WÜRTTEMBERG 2015). Eine potenzielle Ursache für diesen rasanten Anstieg kann in der Win-win-Situation zwischen den Studierenden und den Ausbildungsstätten (= duale Partner) gesehen werden.

Die Ausbildungsstätten profitieren hierbei in mehrfacher Hinsicht von dem dualen Studium und den dual Studierenden (vgl. HOCHSCHULE DUAL 2017). Beispielsweise leisten die Studierenden aktive Arbeit in den Unternehmen, entwerfen Vorschläge für aktuelle Arbeitspakete und -projekte und tragen wesentlich zur Weiterentwicklung des Ausbildungsbetriebs bei. Nach dem Studium stehen den Ausbildungsstätten potenziell wissenschaftlich hochqualifizierte Absolventinnen und Absolventen mit Praxiserfahrung zur Verfügung. Sofern die Betriebe die Ausgebildeten halten können, entfallen Kosten für aufwendige Traineeprogramme, und die Einarbeitungszeit wird bei einer Übernahme auf ein Minimum reduziert (vgl.

DEUER/DUDEK/WINTERHOLLER 2015). Des Weiteren sind die Ausbildungsstätten in der Curriculumgestaltung paritätisch vertreten (vgl. WOLTER 2016; DEUER/TRÄGER 2015).

Auch für Studierende existieren Vorteile durch ein duales Studium. Es dominieren drei Motive zur Aufnahme eines dualen Studiums: Erstes Motiv ist die optimale Verbindung von Theorie und Praxis. Zweites Motiv sind die guten Chancen auf dem Arbeitsmarkt oder sogar eine Übernahme nach dem Studium durch den Ausbildungsbetrieb. Drittes Motiv ist, dass die Studierenden während des Studiums ein eigenes Einkommen haben, das vom Betrieb durchgängig gezahlt wird (vgl. WOLTER u. a. 2014; GENSCHE 2014).

Allerdings müssen die dual Studierenden als eine Art Randgruppe gesehen werden. Erstens werden die Studierenden durch die Einstellung an Ausbildungsstätten stark selektiert, was Voraussetzung für ein duales Studium ist (vgl. KUPFER 2013). Zweitens wurde basierend auf einer vergleichenden Studie der Universität Tübingen gezeigt, dass im Durchschnitt an den Berufsakademien/der DHBW Jugendliche mit besonders guten Noten studieren (vgl. KRAMER u. a. 2011) und sich diese Studienform für Abiturientinnen und Abiturienten mit weniger günstigem familiären Hintergrund und dem Wunsch, rasch finanziell unabhängig zu sein, sowie für Absolventinnen und Absolventen beruflicher Gymnasien als besonders attraktive Alternative zur Universität erweist (vgl. TRAUTWEIN u. a. 2006). Drittens verfügen die Studierenden bereits über berufspraktische Erfahrungen. Folglich können die Studierenden die Arbeitgeberattraktivität in einem gewissen Umfang bewerten und haben Erfahrung mit dieser Thematik. Viertens wird ein duales Studium in verschiedensten Disziplinen angeboten, sodass davon ausgegangen werden kann, dass die Studierenden ihre Studienentscheidung nicht nur nach Studienform, sondern auch nach ihrer Neigung getroffen haben. Innerhalb von sechs Semestern werden im Bachelorstudium an der DHBW beispielsweise aufgrund des Intensivcharakters 210 ECTS-Punkte statt der ansonsten üblichen 180 ECTS-Punkte erworben. Die Studierenden sind zudem in Ausbildungsstätten durchgängig beschäftigt und erhalten von dort ein fortlaufendes Gehalt. Folglich unterscheiden sich die Verhältnisse an der DHBW von den bereits stark beforschten Lernumwelten der Universitäten oder Fachhochschulen und bilden einen besonderen Lehr- und Lernort für die Akteurinnen und Akteure in den Ausbildungsstätten sowie an der Hochschule (vgl. DEUER/TRÄGER 2015).

2 Datenerhebung

Soll die Wirkung einer Intervention im Rahmen akademischer Forschung abgeschätzt werden, dann stellt sich in der Regel das Problem, dass kaum auf die als „Goldstandards“ bezeichneten *Random Controlled Trials* zurückgegriffen werden kann. *Matching*-Verfahren bieten in der Bildungsforschung eine Möglichkeit, auch für nichtexperimentelle Studien kausale Inferenzen zu rechtfertigen (vgl. BECKER 2011). Die vorliegende Studie untersucht unter der Zuhilfenahme des in der Praxis gut etablierten *Propensity Score Matchings*, wie sich anhand nichtparametrischer Verfahren kausale Effekte in Beobachtungsdaten statistisch abschätzen lassen. Bei diesem Vorgehen basiert die Drittvariablenkontrolle auf der Konstruk-

tion statistischer Vergleichsgruppen, die in Bezug auf relevante Hintergrundfaktoren weitgehend identisch sind (vgl. GANGL 2010). Wir verwenden diese Methode, um zu untersuchen, ob Studierende abhängig von der Schulzeit bis zum Abitur unterschiedliche Auffassungen von Arbeitgeberattraktivität entwickeln.

2.1 Stichprobe

Die Basis der empirischen Analyse bildet eine Stichprobe von $n = 299$ Datensätzen, mit einer effektiven Stichprobe von $n = 276$ bei den multivariaten Analysen (u. a. wurden Ausreißer entfernt). Die Datenerhebung erfolgte als *Paper-and-Pen*-Umfrage im Querschnittsdesign unter Studierenden der DHWB Ravensburg von Mai bis Juli 2016. Die Stichprobenauswahl erfolgte anhand einer Klumpenstichprobe und wurde realisiert, indem die Studiengangsleitungen gebeten wurden, die Fragebögen in Studienkohorten vollständig zu verteilen. Das Durchschnittsalter der Befragten beträgt 22 Jahre ($M = 21,68$, $SD = 2,04$). Die Hälfte der Studierenden (49,16 %) absolvierte das Abitur in der achtjährigen Sekundarstufenzeit, 50,84 Prozent in der neunjährigen Sekundarstufenzeit. Insgesamt gaben 81,61 Prozent der Befragten an, das Abitur an einem baden-württembergischen Gymnasium abgelegt zu haben. Am stärksten vertreten ist der Abiturjahrgang 2013, die Spanne reicht jedoch insgesamt von Jahrgang 2005 bis Jahrgang 2015. Die Studierenden verteilen sich auf vier Bachelorstudiengänge: BWL-Industrie (52,84 %), Wirtschaftsinformatik (11,37 %), Maschinenbau (30,77 %) und Wirtschaftsingenieurwesen (5,02 %). Der Schwerpunkt auf diesen Studienfächern ist damit zu begründen, dass in Deutschland in diesen Studienrichtungen die meisten Studierenden eingeschrieben sind (DESTATIS 2019). 39,13 Prozent der Studierenden gaben an, im ersten Studienjahr, 30,1 Prozent gaben an, im zweiten Studienjahr und 30,77 Prozent gaben an, im dritten Studienjahr zu sein. Ein Teil der Studierenden hatte bereits vor Aufnahme des Studiums eine Ausbildung absolviert (13,09 %).

2.2 Erhebungsinstrumente

Die Beurteilung der Arbeitgeberattraktivität wurde über vier Skalen erhoben, die unterschiedlichen, bereits existierenden Instrumenten entnommen wurden. Die Skala „Sicherheit und Perspektive“ sowie „Sinnerfüllung“ wurde Leven u. a. (2015) entnommen, die Skala „Innovationspotenzial“ modifizierten wir gemäß Ramm u. a. (2014), ebenso die Skala „Work-Life-Orientierung“, die wir in Anlehnung an Leven u. a. (2015) auch in einer modifizierten Form einsetzten. Die Bewertung der Fragen erfolgte über eine fünfstufige Likert-Skala, deren Skalenendpunkte mit 1 (= trifft nicht zu) und 5 (= trifft voll zu) verbal verankert wurden.

Die Reliabilität der Skalen ermittelten wir durch eine vorgeschaltete und hier nicht abgebildete Faktorenanalyse, für die insgesamt 299 Datensätze zur Verfügung standen. In Tabelle 1 sind die Kennwerte der Skalen zusammengefasst. Die Reliabilitäten der Skalen „Sinnerfüllung“ und „Work-Life-Orientierung“ betragen Cronbachs $\alpha_c = .81$ und $c_\alpha = .76$.

Auch die Werte der Skala „Innovationspotenzial“, $c_\alpha = .69$, sowie der Skala „Sicherheit und Perspektive“, $c_\alpha = .68$, liegen innerhalb des akzeptablen Bereichs.

Tabelle 1: Skalenkennwerte der psychometrischen Messinstrumente von Arbeitgeberattraktivität

	<i>M</i>	<i>SD</i>	Itemzahl	c_α	Itembeispiel: Was müsste Ihnen eine berufliche Tätigkeit bieten, damit Sie zufrieden sein können:
Sicherheit und Perspektive	4,26	0,66	3	0,68	... ein hohes Einkommen.
Sinnerfüllung	4,27	0,64	4	0,81	... das Gefühl, etwas zu leisten.
Innovationspotenzial	2,84	0,90	2	0,69	... die Möglichkeit, Unbekanntes zu erforschen.
<i>Work-Life-Orientierung</i>	4,29	0,73	4	0,76	... genügend Freizeit neben der Berufstätigkeit.

Anmerkung: $n = 299$; Likert-Skala von 1 (= trifft nicht zu) bis 5 (= trifft voll zu)

Ein Blick auf die Mittelwerte der Skalen zeigt, dass zunächst allen Kriterien eine gewisse Wichtigkeit beigemessen wird. Drei von vier Skalen liegen durchschnittlich sogar über vier Punkten („Sicherheit und Perspektive“, „Sinnerfüllung“ sowie „*Work-Life-Orientierung*“). Allein diese Information legt nahe, dass diese drei Aspekte demnach für die untersuchte Generation Merkmale attraktiver Betriebe sind. Der Mittelwert für das Potenzial für Innovationen scheint weniger wichtig zu sein – er beträgt $M = 2,84$. Hier gilt es allerdings die große Standardabweichung von $SD = 0,9$ zu berücksichtigen.

Die Auswahl der Kovariaten wurde wie folgt vorgenommen: Das Jahr der Hochschulzugangsberechtigung ist wichtig, da sich die Aufgaben für das Abitur jedes Jahr verändern. Basierend auf der Überlegung, dass das soziale und kulturelle Kapital mit jedem Jahr an der Hochschule wächst, wird das Studienjahr ebenfalls berücksichtigt (vgl. BOURDIEU 1983; ENGLER 1993; VOSGERAU 2005). Aufgrund der föderalen Struktur im Bildungswesen wurde die Variable Bundesland einbezogen. Und basierend auf Leistungsunterschieden im deutschen Schulsystem, die auf das Geschlecht zurückgeführt werden, wurde diese Variable ebenfalls berücksichtigt (vgl. HURRELMANN/SCHULTZ 2012). Der Studiengang wurde als Kovariate integriert, da sich die Studierenden in diesen Merkmalen aufgrund der Fachsozialisation zum Teil stark unterscheiden können (vgl. ENGLER 1993; VOSGERAU 2005). Da es fragwürdig ist, ob komplexere Modelle mit mehr Kovariaten bessere Ergebnisse hervorbringen, wurden keine weiteren Variablen in die Analyse integriert (vgl. BRÜDERL 2010).

2.3 Datenanalyse

Grundsätzlich ist zu bemerken, dass an eine Schätzung kausaler Wirkungen sozialer Interventionen hohe Anforderungen bestehen müssen. Im Rahmen kausaler Wirkungsevaluationen wird aktuell ein Diskurs gepflegt, der eben diese Thematik aufgreift (vgl. MÜLLER 2017). Im Kern geht es um die möglichst fehlerfreie Zuschreibung einer Wirkung zu einer

bestimmten Intervention. Da auf die als „Goldstandard“ charakterisierten Designs der *Random Controlled Trials* (RCTs) meistens nicht zurückgegriffen werden kann, beispielsweise aus ethischen Gründen bei der Gruppenzuteilung, werden Alternativen eingesetzt. *Matching*-Verfahren bieten hierbei eine immer populärer werdende Möglichkeit in der Bildungsforschung (vgl. BECKER 2011), um Drittvariablen zu kontrollieren. Somit besteht in dieser Vorgehensweise eine Möglichkeit, der bewussten oder unbewussten Einflussnahme beim Identifizieren kausaler Zusammenhänge in der quantitativen Datenanalyse (kausale Inferenz) eine methodische Maßnahme entgegenzusetzen.

Ein ideales Vorgehen für einen validen Vergleich zwischen, im vorliegenden Fall, den beiden Beschulungsdauern G8 und G9 bestünde in einer randomisierten Zuweisung der Schülerinnen und Schüler zu einer Schulart, wie es beispielsweise bei einem experimentellen Versuchsplan erfolgen würde und in der Wirkungsevaluationsforschung oft thematisiert wird (vgl. MÜLLER 2017). Im vorliegenden Fall ist dies nicht möglich. Folglich können Unterschiede in den betrachteten Gruppen bestehen, die bereits vor Studiendurchführung oder durch andere Selektionsverfahren zustande gekommen sind (vgl. MORGAN/WINSHIP 2007; PAN/BAI 2015).

Eine Möglichkeit, diesen Bias statistisch zu kontrollieren, ist das *Propensity Score Matching*. Durch dieses Verfahren werden mögliche Gruppenunterschiede auf der Basis theoretisch hergeleiteter Faktoren beseitigt, um den Beschulungseffekt genau identifizieren zu können (vgl. ROSENBAUM/RUBIN 1983). Hierbei berechnet das *Propensity Score Matching* für jeden Datensatz der Stichprobe auf Grundlage der theoretisch abgeleiteten relevanten Merkmale den Wahrscheinlichkeitswert, den sogenannten *Propensity Score*, für den Einfluss der Beschulungsdauer auf die Beurteilung der Merkmale von Arbeitgeberattraktivität. Im Anschluss werden dann nur diejenigen Datensätze beider Gruppen (G8/G9) in ihren Merkmalsausprägungen verglichen, die ähnliche *Propensity Scores* besitzen. Auf der Grundlage der nachstehenden fünf Schritte wird das *Propensity Score Matching* in dieser Untersuchung angewendet (vgl. BECKER 2011).

Der erste Schritt beinhaltet eine Prüfung durch *t*-Tests und χ^2 -Tests, ob ein systematischer Unterschied oder Zusammenhang zwischen der Beschulungsdauer (G8/G9) und anderen wirkenden Merkmalen (Hintergrundmerkmale/Kovariaten) besteht. Im darauffolgenden zweiten Schritt analysieren wir auf Grundlage der definierten Kontrollvariablen den *Propensity Score* für die Regelschulzeit bis zum Abitur anhand einer logistischen Regression. Der dritte Schritt beinhaltet das *Matching*-Verfahren. Wir verwendeten drei Verfahren:

1. das *Nearest-Neighbour-Matching* (NN 1), bei dem eine Paarbildung zwischen einer Beobachtung von G8 und der ähnlichsten Beobachtung von G9 vorgenommen wurde,
2. das *Nearest-Neighbour-Matching* (NN 5), bei dem eine Paarbildung zwischen einer Beobachtung der G8-Kohorte und den fünf ähnlichsten Beobachtungen von G9-Kohorten vorgenommen wurde, und

3. das *Kernel-Matching* (*Kernel*), bei dem eine Paarbildung zwischen einer Beobachtung einer G8-Kohorte und der gesamten G9-Kohorte als gewichteter Durchschnitt vorgenommen wird (GANGL 2010).

In einem vierten Schritt untersuchen wir, ob die zwei Gruppen vergleichbar und die *Matching*-Verfahren gelungen sind. Wir ziehen hierzu die standardisierten Differenzen heran, um die jeweils neue Balance zu überprüfen. Je geringer die standardisierten Differenzen nach dem *Matching* ausfallen, desto mehr gleichen sich diese Untergruppen. Nach Sicherstellung einer hinreichenden Ähnlichkeit wird der fünfte Schritt, das inferenzstatistische Auswertungsverfahren mit der gematchten Stichprobe, durchgeführt. Im Anschluss wird der *Average Treatment Effect on the Treated* (*ATT*) berechnet (vgl. BECKER 2011). Dieser Effekt fungiert als Index, der die Unterschiede in der durchschnittlichen Merkmalsausprägung zwischen der G8- und G9-Kohorte zeigt. Zur Validierung der Methode wird abschließend eine Sensitivitätsanalyse nach dem Ansatz des *Rosenbaum bounds* (vgl. ROSENBAUM 2002) durchgeführt. Wir richten uns nach dem Kriterium von Aakvik (2001), das bei einem Γ von 2.0 mit $p < .1$ (*upper bound significance level*) ein robustes Ergebnis ansetzt. Mit anderen Worten, ist das Aakvik-Kriterium erfüllt, können wir davon ausgehen, dass die integrierten Kovariaten die errechneten Effekte tatsächlich beeinflussen. Für die Datenanalyse verwendeten wir die Statistikprogramme STATA (Version 14) und SPSS (Version 24). Anhand des ado `psmatch2` berechneten wir die Ergebnisse des *Propensity Score Matchings*.

3 Ergebnisse

Die Verteilung der Hintergrundkovariaten in den Gruppen G8-Abitur und G9-Abitur ist Tabelle 2 zu entnehmen (Schritt 1). Alle durchgeführten Zusammenhangs- und Unterschiedsanalysen ergaben signifikante Unterschiede und bestätigen, dass ein *Propensity Score Matching* anwendbar ist. Der deutlichste Zusammenhang zeigte sich für die Variable Bundesland [$\chi^2(1,299) = 28,755; p < 0,001; \Phi = 0,31$]. Aber auch der Studiengang [Studiengang Technik der G8-Kohorte, $\chi^2(1,299) = 8,701; p = 0,003; \Phi = 0,171$] und das Geschlecht [Studentinnen der G9-Kohorte, $\chi^2(1,299) = 4,745; p < 0,029; \Phi = 0,126$] zeigten signifikante Unterschiede. Die Analysen weisen darüber hinaus nach, dass Studierende der G8-Kohorte ($M = 1,73; SD = 0,80$) in einem niedrigeren Studienjahr sind als Studierende der G9-Kohorte ($M = 2,09; SD = 0,83$). Das Ergebnis ist signifikant [$t(297) = 3,790; p < 0,001; d = 0,44$]. Das Jahr der Hochschulzugangsberechtigung war zwischen G8 ($M = 2013,87; SD = 1,05$) und G9 ($M = 2013,11; SD = 1,95$) ebenfalls signifikant unterschiedlich [$t(233,22) = 4,246; p < 0,001; d = 0,49$].

Tabelle 2: Verteilung der Hintergrundkovariaten in den Vergleichsgruppen

Merkmalsausprägung	G8-Abitur (Anteil in % / Mittelwert)	G9-Abitur (Anteil in % / Mittelwert)
Studengang		
Technik	45,58	26,32
Wirtschaftswissenschaft	54,21	73,68
Bundesland		
Bayern	30,61	06,58
Baden-Württemberg	69,39	93,32
Geschlecht		
männlich	61,22	48,68
weiblich	38,78	51,32
Studienjahr	01,73	02,09
Abiturjahr	2013,87	2013,11
n	147,00	152,00

Anmerkung: n = 299

Durch eine logistische Regressionsanalyse überprüften wir dann, welche Hintergrundkovariaten Prädiktoren für die Beschulung in einem G8- oder G9-Zug sind (Schritt 2). Die in Tabelle 3 zusammengefassten Ergebnisse korrespondieren im Wesentlichen mit den deskriptiven Befunden; Studierende der Wirtschaftswissenschaft, die in Baden-Württemberg ihre Hochschulreife erworben haben und dies bereits vor einigen Jahren taten, sind beispielsweise typischerweise Absolvierende eines Gymnasiums mit neun Schuljahren in der Sekundarstufe 2.

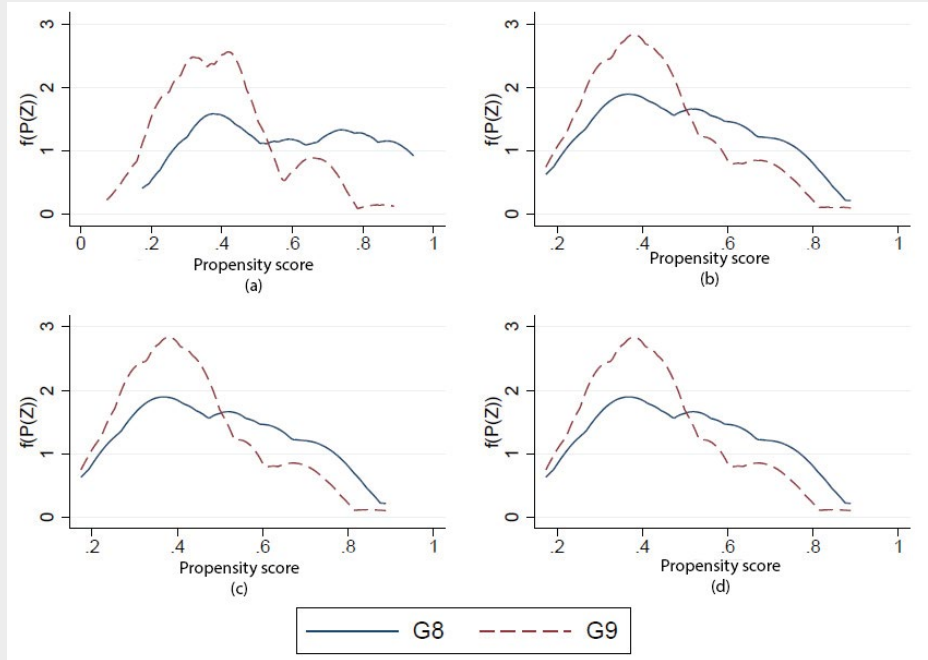
Tabelle 3: Ergebnisse der logistischen Regression zur Vorhersage der Vergleichsgruppen

	<i>B</i>	<i>SE</i>	<i>p</i>
Studengang	-0,74	0,30	0,014
Bundesland	-2,08	0,41	≤ 0,001
Geschlecht	0,13	0,29	0,628
Studienjahr	-0,15	0,19	0,431
Abiturjahr	0,28	0,10	0,007
R ² (Nagelkerke)	14,83		

Anmerkung: n = 282

Der darauffolgende dritte Schritt beinhaltet die Durchführung des *Matchings*, die Zuordnung gleicher Fälle. Wir führten die *Matching*-Verfahren NN 1, NN 5 und *Kernel* durch, um die Sensibilität der Lösungen zu ermitteln (vgl. BECKER 2011).

Abbildung 1: Kerndichteschätzer der Verteilungsfunktion von den *Propensity scores* (n = 276 bzw. n = 299) vor und nach dem *Matching*

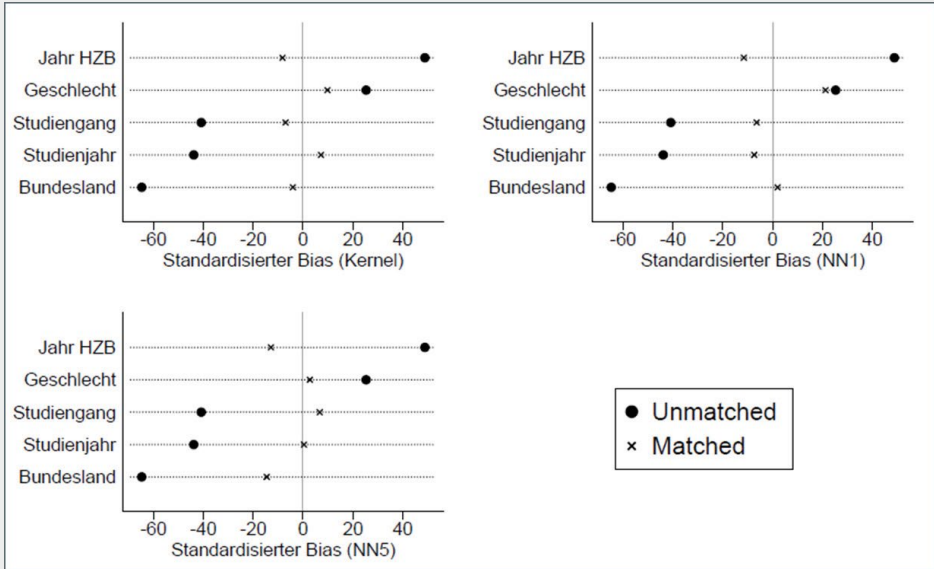


Anmerkung: (a) Rohwerte; (b) *Matching* nach *Kernel*, (c) *Matching* nach *Nearest Neighbor 1*; (d) *Matching* nach *Nearest Neighbor 5*.

Quelle: eigene Darstellung

Die Verteilung der *Propensity Scores* und die Bewertung der *Matching*-Verfahren anhand des standardisierten Bias sind Gegenstand des vierten Analyseschritts, der Evaluation des *Matchings*. Anhand der Verteilungskurven in Abbildung 1 lässt sich erkennen, dass sich die berechneten *Propensity Scores* aufgrund ihrer zum Teil großen Schnittmenge für ein *Matching* eignen. Die aus Abbildung 2 zu entnehmenden abgebildeten standardisierten Mittelwertdifferenzen zeigen, dass sich der *Bias* des Abiturjahres (Jahr der Hochschulzugangsberechtigung, HZB) und der des Bundeslands der HZB vor und nach dem *Matching* stark reduziert haben. Lediglich die geringe Zahl von vier Kovariaten liegt über der Zehnprozenthürde für standardisierte Mittelwertdifferenzen nach dem *Matching*.

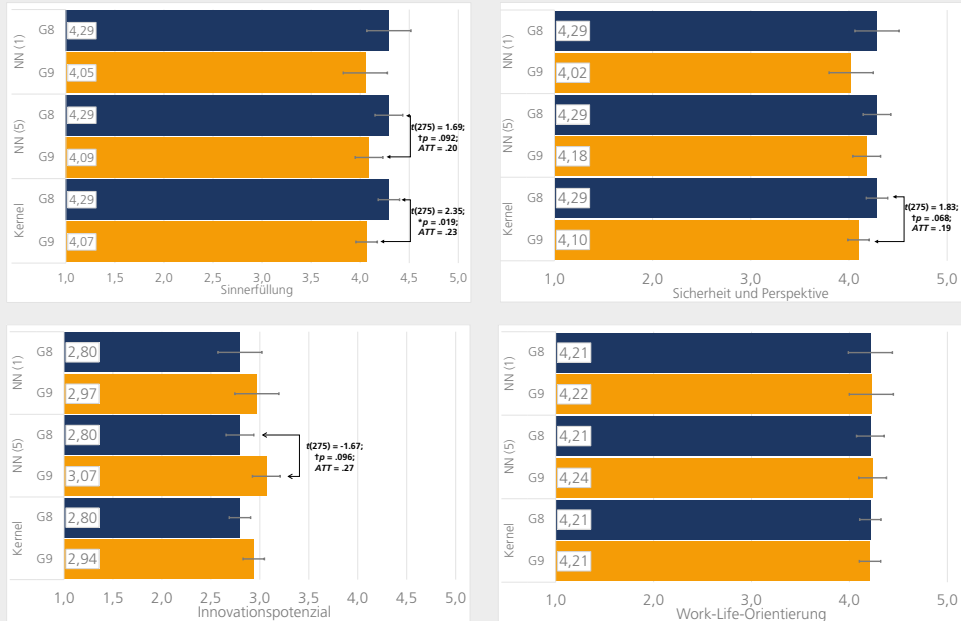
Abbildung 2: Standardisierte Mittelwertdifferenzen zwischen der Beschulung G8 und G9 vor und nach dem Matching



Quelle: eigene Darstellung

Der fünfte und letzte Analyseschritt beinhaltet die Schätzung der Effektstärke für die Beschulungen nach G8 und G9 im Hinblick auf die Wahrnehmung der Arbeitgeberattraktivität. Die in Abbildung 3 dargestellten Ergebnisse zeigen folgende Trends: Es kann davon ausgegangen werden, dass Abiturientinnen und Abiturienten im G8-Zug eine sinnerfüllendere Arbeit ausüben wollen als G9-Abiturientinnen und -Abiturienten. Dies kann anhand der (marginal) signifikanten *ATTs* um 0,2 bei den *Matching*-Verfahren *Kernel* und *Nearest-Neighbor 5* belegt werden. Des Weiteren streben G8-Abiturientinnen und -Abiturienten stärker nach Sicherheit und Perspektive in einer Erwerbsarbeit als Abiturientinnen und Abiturienten des G9. Der *Matching*-Algorithmus *Kernel* weist hier ein Ergebnis von $ATT = 0,19$ auf. Dagegen zeigt sich, dass G9-Abiturientinnen und -Abiturienten die Förderung des individuellen Innovationspotenzials signifikant wichtiger finden als die Abiturientinnen und Abiturienten der G8-Kohorte bei $ATT = 0,27$ für *Nearest-Neighbor 5*. Keine signifikanten Unterschiede bestehen in der *Work-Life*-Orientierung.

Abbildung 3: Kennwerte (u. a. CI 95 %) und Effekte nach Beschulungsart für die Merkmale von Arbeitgeberattraktivität nach dem *Matching*



Anmerkung: ATT = Average treatment effect on the treated

Quelle: eigene Darstellung

Inwiefern die ausgewählten Kovariaten relevante Einflussgrößen sind, untersuchten wir anhand einer Sensitivitätsanalyse. Wir entschieden uns für die *Rosenbaum-bounds*-Methode und das strenge Aakvik-Kriterium. Die Ergebnisse sind in Tabelle 4 aufgeführt. Jedem *Matching* wurde jede Form der gemessenen Arbeitgeberattraktivität gegenübergestellt. Für das Merkmal „Sicherheit und Perspektive“ zeigt die Analyse, dass die angesetzten Grenzwerte von $\Gamma = 2,0$ mit $p < 0,1$ für *Nearest Neighbor 1* ($\Gamma = 2,0$ mit $p = 0,17$) und *Kernel* ($\Gamma = 2,0$ mit $p = 0,18$) nur knapp verfehlt wurden. Ähnlich ist das Ergebnis für das Attraktivitätsmerkmal „Sinnerfüllung“. Hier scheitert ein positives Ergebnis für die *Matching*-Verfahren *Kernel* ($\Gamma = 2,0$ mit $p = 0,13$) und *Nearest Neighbor 5* ($\Gamma = 2,0$ mit $p = 0,13$) am strengen Grenzwertkriterium minimal. Die Sensitivitätsanalyse für das Innovationspotenzial zeigt, dass alle *Matching*-Verfahren unter den Grenzwerten bleiben und somit von einer robusten Schätzung auszugehen ist. Für die Analyse der *Work-Life-Orientierung* sind die Ergebnisse schwer zu interpretieren, da die Analysemuster gegensätzliche Messwerte für die *Matching*-Verfahren *Nearest Neighbor 5* und *Kernel* ausgeben bzw. der Grenzwert für *Nearest Neighbor 1* von $p < 0,01$ nicht mal bei $\Gamma = 1,0$ erreicht wird. Offenbar scheinen hier andere Kovariaten

zu wirken, die bei zukünftigen Forschungen in diesem Kontext eingesetzt werden sollten. Insgesamt zeigt die durchgeführte Sensitivitätsanalyse, dass die Ergebnisse weitestgehend robust sind und angebrachte Kovariaten für die Analysen eingesetzt wurden.

Tabelle 4: Sensitivitätsanalyse der geschätzten Propensity Score Matching-Modelle nach Rosenbaum bounds

	Nearest Neighbor 1		Nearest Neighbor 5		Kernel	
	Γ	p*	Γ	p*	Γ	p*
Sicherheit und Perspektive	1,0	<0,001	1,0	<0,001	1,0	<0,001
	1,2	<0,001	1,2	0,010	1,2	<0,001
	1,4	<0,010	1,4	0,060	1,4	<0,010
	1,6	0,020	1,6	0,190	1,6	0,030
	1,8	0,070	1,8	0,370	1,8	0,080
	2,0	0,170	2,0	0,570	2,0	0,180
Sinnerfüllung	1,0	<0,001	1,0	<0,001	1,0	<0,001
	1,2	<0,010	1,2	<0,001	1,2	<0,001
	1,4	0,01	1,4	<0,010	1,4	<0,010
	1,6	0,06	1,6	0,020	1,6	0,020
	1,8	0,15	1,8	0,050	1,8	0,050
	2,0	0,28	2,0	0,130	2,0	0,130
Innovationspotenzial	1,0	<0,010	1,0	<0,010	1,0	0,020
	1,2	<0,001	1,2	<0,001	1,2	<0,010
	1,4	<0,001	1,4	<0,001	1,4	<0,001
	1,6	<0,001	1,6	<0,001	1,6	<0,001
	1,8	<0,001	1,8	<0,001	1,8	<0,001
	2,0	<0,001	2,0	<0,001	2,0	<0,001
Work-Life-Orientierung	1,0	0,12	1,0	0,72	1,0	0,89
	1,2	0,37	1,2	0,37	1,2	0,63
	1,4	0,67	1,4	0,13	1,4	0,33
	1,6	0,86	1,6	0,04	1,6	0,13
	1,8	0,95	1,8	<0,01	1,8	0,04
	2,0	0,98	2,0	<0,01	2,0	0,01

4 Diskussion

Ziel dieser Studie war, den Einfluss der Beschuldungsdauer bis zum Abitur (G8 und G9) auf die Beurteilung der Arbeitgeberattraktivität von dual Studierenden zu untersuchen. Hierzu wurde eine Primärdatenerhebung unter DHBW-Studierenden durchgeführt. Mithilfe des *Propensity Score Matchings* konnte gezeigt werden, dass signifikante Unterschiede in den Merkmalen „Sicherheit und Perspektive“ sowie „Sinnstiftung“ und „Innovationspotenzial“ bestehen.

Vor dem eingangs dargestellten Forschungsstand kann das Ergebnis weiter interpretiert werden. Absolventinnen und Absolventen des G8-Abiturs schätzen die Attraktivitätsmerkmale „Sicherheit und Perspektive“ sowie „Sinnstiftung“ signifikant höher ein als die G9-Kohorte. Folgt man den theoretischen Überlegungen und Studienergebnissen aus Abschnitt 2, so lässt sich dieser Befund schlüssig in den Forschungsstand einordnen: Das achtjährige Gymnasium setzt die Schülerinnen und Schüler einem verdichteten Curriculum sowie erhöhtem zeitlichen Leistungsdruck aus. Infolgedessen bleibt weniger Zeit, das eigene Innovationspotenzial zu explorieren oder zu verwirklichen. Diesen Schluss unterstützt der Befund, dass G8-Absolvierende weniger Wert auf Innovationsförderung im Betrieb legen sowie ein geringeres Interesse an Innovationsförderung haben, aber eine starke Orientierung an sinnerfüllender Arbeit zeigen. Diese beiden Aspekte in Kombination können auch ein Hinweis darauf sein, dass das achtjährige Gymnasium Studierende hervorbringt, die stärker auf der Suche nach der eigenen Identität sind. Eventuell wird unter der Annahme einer hohen Belastung und geringeren Freizeit in G8-Jahrgängen diese Entwicklung gehemmt. Beide Studierendengruppen wünschen sich eine ausgewogene *Work-Life-Orientierung*. Das Freizeitdefizit-Argument der G8-Kritikerinnen und -Kritiker ließe sich wohl nur untermauern, wenn die G8-Stichprobe eine signifikant höhere Ausprägung dieser Dimension hätte.

Eine weitere Interpretationsmöglichkeit der Bewertungsdifferenzen zwischen G8- und G9-Absolventinnen und -Absolventen ist, dass Studierende aus dem achtjährigen Gymnasium ein zügigeres Lerntempo gewohnt sind und zielstrebig ein duales Studium absolvieren, um bestmögliche Aussichten auf eine Stelle in der Nähe ihres Wohnorts zu haben. Für diese Interpretation spricht zum einen, dass die deutliche Mehrheit der dual Studierenden an der DHBW aus Baden-Württemberg stammt, zum anderen ist dies kongruent mit den eingangs erläuterten Ergebnissen der großen Jugendstudien. Insgesamt konnte also gezeigt werden, dass die Regelschulzeit bis zum Abitur (G8/G9) unter Studierenden der DHBW zu unterschiedlichen Vorstellungen von Arbeitgeberattraktivität sozialisiert. Folgestudien an anderen Hochschulen sind wünschenswert, um eine hochschulunabhängige Aussage treffen zu können.

Was bedeuten die Studienbefunde für Unternehmen? Die vorliegende Untersuchung zeigt, dass jede Generation andere Erwartungen an ihren Arbeitgeber hat als die vorherige Generation. Möchten Betriebe qualifizierte Mitarbeiterinnen und Mitarbeiter gewinnen und diese nicht an die Konkurrenz verlieren, sollten sie sich überlegen, inwiefern sie auf Befunde

dieser und ähnlicher Studien reagieren können. Verlangen beispielsweise Absolventinnen und Absolventen des achtjährigen Gymnasiums vor allem, einen sinnstiftenden Arbeitsplatz ausfüllen zu dürfen, könnten Unternehmen neue Mitsprachestrukturen implementieren.

Was bedeuten die Studienergebnisse für Hochschulen? Durch die Verkürzung der Schulzeit müssen sich die Hochschulen auf jüngere Studienanfängerinnen und Studienanfänger einstellen, die über weniger Lebenserfahrung verfügen. Des Weiteren wird durch die Differenzierung der Schulformen und landesabhängigen Schulreformen die Heterogenisierung und Ausdifferenzierung der Studierendenschaft weiter verstärkt werden (vgl. DOBISCHAT/DÜSSELDORFF 2015). Hochschulen mit unmittelbar berufsqualifizierenden Studienabschlüssen müssen sich fragen, wie sie die Thematik der Anforderungen an den Arbeitsplatz aufgreifen, da es offensichtlich unterschiedliche Prioritäten bei Personengruppen gibt. Insbesondere hat der Wissenschaftsrat (2015) als wesentliche Aufgaben der Hochschulen – neben der fachwissenschaftlichen Qualifizierung – die Förderung der Persönlichkeitsbildung und die Arbeitsmarktvorbereitung identifiziert.

Einschränkungen der Ergebnisinterpretation beruhen teils auf geringen Effektstärken, die auf die unterschiedlichen *Matching*-Algorithmen zurückzuführen sind. Des Weiteren ist zu beachten, dass Langzeiteffekte nicht berücksichtigt werden konnten und insofern kausale Interpretationen eingeschränkte Gültigkeit haben. Es wurden eventuell relevante Kontrollvariablen nicht erhoben. Dies gilt insbesondere für Klassenwiederholungen, die in der Befragung nicht explizit erhoben worden sind, aber einen Einfluss auf das Ergebnis haben könnten. Folgestudien könnten beispielsweise in der Erweiterung der Stichprobe um andere Bundesländer oder Studiengänge bestehen. Dies würde dann eine Auswertung der Daten mit einer Mehrebenenanalyse ermöglichen und komplexe Datenstrukturen offenlegen. Auf diese Weise könnte nach und nach ein systematischer Forschungsstand zu Auswirkungen der Regelschulzeit bis zum Abitur aufgebaut werden. Allerdings muss bei der vorliegenden Studie konstatiert werden, dass das methodische Vorgehen und die Datenanalyse unter den Ansätzen einer wirkungsevaluationstheoretischen Herleitung und Betrachtungsweise vorgenommen wurden. Dies steigert die Aussagekraft der Studie. Parallel dazu stehen zukünftige Studien in diesem Forschungsbereich vor der Aufgabe, die nachwachsende Generation Z in den Fokus zu nehmen, da diese Personen immer mehr an die Hochschulen drängen (vgl. SCHOLZ 2014; STECKL/SIMSHÄUSER/NIEDERBERGER 2019; ANDRIONE 2018). Hierbei gilt es verstärkt auch aktuell wichtige Aspekte wie Bildung für eine nachhaltige Entwicklung, digitale Kompetenzen (vgl. SCHULZE HEULING/WILD 2019; WILD/SCHULZE HEULING 2020) sowie Gesundheitsförderung in die Forschungen zu integrieren. Basierend auf derartigen Grundlagen sind gehaltvolle Studien zu erwarten.

Literatur

- AAKVIK, Arild: Bounding a matching estimator: The case of a Norwegian training program. In: Oxford Bulletin of Economics and Statistics (2001) 63, S. 115–143
- ALDERFER, Clayton P.: Existence, Relatedness and Growth. Human Needs in Organizational Settings. New York 1972
- ANDRIONE, Ludwig: Eigenschaften, Einstellungen und Werte von Generationen: Stand und Aussicht der Forschung. In: Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO) (2018) 49, S. 415–419
- AUSBILDUNGPLUS: Duales Studium in Zahlen. Trends und Analysen. 2014 – URL: https://www.bibb.de/dokumente/pdf/duales-studium_in_zahlen_2014.pdf (Zugriff: 05.04.2017)
- AUTORENGRUPPE BILDUNGSBERICHTERSTATTUNG: Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. Bielefeld 2016
- BAYERISCHES STAATSMINISTERIUM FÜR BILDUNG UND KULTUS, WISSENSCHAFT UND KUNST (Hrsg.): Das neue bayerische Gymnasium. So entwickelt sich das Gymnasium zeitgemäß weiter – URL: <https://stmwk.bayern.de/allgemein/meldung/5082/so-entwickelt-sich-das-gymnasium-zeitgemaess-weiter.html> (Zugriff: 09.03.2021)
- BECKER, Michael: Matching-Verfahren und Gruppenvergleiche. In: MASCHKE, Sabine; STECHER, Ludwig (Hrsg.): Enzyklopädie Erziehungswissenschaft Online. Fachgebiet: Methoden der empirischen erziehungswissenschaftlichen Forschung, Quantitative Methoden. Weinheim 2011, S. 1–50
- BOURDIEU, Pierre: Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In: KRECKEL, Reinhard (Hrsg.): Soziale Ungleichheiten. Soziale Welt Sonderband 2. Göttingen 1983, S. 183–198
- BRÜDERL, Josef: Kausalanalysen mit Paneldaten. In: WOLF, Christof; BEST, Henning (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden 2010, S. 963–994
- BURCHARD, Amory: Streit ums Abitur. Bildungsforscher fürchten den Elternwillen. Der Tagesspiegel vom 22.08.2016 – URL: <http://www.tagesspiegel.de/wissen/streit-ums-abitur-bildungsforscher-fuerchten-den-elternwillen/14436110.html> (Zugriff: 09.09.2016)
- BÜTTNER, Bettina; THOMSEN, Stephan L.: Are we spending too many years in school? Causal evidence of the impact of shortening secondary school duration. In: German Economic Review (2015) 16, S. 65–86
- DAHMAN, Sarah; ANGER, Silke: The impact of education on personality: Evidence from a German high school reform – URL: https://www.diw.de/documents/publikationen/73/diw_01.c.465765.de/diw_sp0658.pdf (Zugriff: 10.10.2016)
- DERR, Katja; HÜBL, Reinhold; ZAKI, Ahmed: Online-Eingangstests und Lernmaterialien zur Studienvorbereitung Mathematik in den Ingenieurwissenschaften. In: HOPPENBROCK, Axel u. a. (Hrsg.): Mathematik im Übergang Schule/Hochschule und im ersten Studienjahr. Kassel 2013, S. 35–36

- DESTATIS (Hrsg.): Studierende nach Fächergruppen. Insgesamt nach Fächergruppen – URL: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Tabellen/studierende-insgesamt-faechergruppe.html> (Zugriff: 09.08.2019)
- DEUER, Ernst; DUDEK, Heinz-Leo; WINTERHOLLER, Maria: Standortfaktor Duale Hochschule – Perspektiven für Jugendliche, Fachkräftepotenziale für die Betriebe und Kaufkraft für die Region. Ravensburg 2015
- DEUER, Ernst; TRÄGER, Manfred: Duale Studiengänge eröffnen neue Potenziale für die betriebliche Personalarbeit. In: BRÜGGEMANN, Tim; DEUER, Ernst (Hrsg.): Berufsorientierung aus Unternehmenssicht. Fachkräfterekrutierung am Übergang Schule – Beruf. Bielefeld 2015, S. 187–198
- DEUTSCHE GESELLSCHAFT FÜR PERSONALFÜHRUNG (Hrsg.): Zwischen Anspruch und Wirklichkeit: Generation Y finden, fördern und binden – URL: <https://static.dgfp.de/assets/publikationen/2011/GenerationY-findern-foerdern-bindern.pdf> (Zugriff: 17.09.2016)
- DOBISCHAT, Rolf; DÜSSELDORFF, Karl: Sozialisation in Berufsbildung und Hochschule. In: HURRELMANN, Klaus u. a. (Hrsg.): Handbuch Sozialisationsforschung. 8. Aufl. Weinheim 2015, S. 469–491
- ENGLER, Steffanie: Fachkultur, Geschlecht und soziale Reproduktion: eine Untersuchung über Studentinnen und Studenten der Erziehungswissenschaft, Rechtswissenschaft, Elektrotechnik und des Maschinenbaus. Weinheim 1993
- GANGL, Markus: Nichtparametrische Schätzung kausaler Effekte mittels Matchingverfahren. In: WOLF, Christof; BEST, Henning (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden 2010, S. 931–961
- GENSCH, Kristina: Dual Studierende in Bayern – Sozioökonomische Merkmale, Zufriedenheit, Perspektiven – URL: http://www.ihf.bayern.de/uploads/media/IHF_Studien_zur_Hochschulforschung-84.pdf (Zugriff: 09.03.2017)
- HOCHSCHULE DUAL (Hrsg.): Heute die Besten für morgen gewinnen – URL: <http://www.hochschule-dual.de/unternehmen/vorteile-des-dualen-studiums/darum-dual-zu-personal.html> (Zugriff: 01.04.2017)
- HOFMANN, Silvia u. a.: AusbildungPlus. Duales Studium in Zahlen 2019. Trends und Analysen. Bonn 2019 – URL: https://www.bibb.de/dokumente/pdf/06072020_AiZ_dualesStudium-2019.pdf (Zugriff: 25.08.2020)
- HOLSTE, Jan Hauke: Arbeitgeberattraktivität im demographischen Wandel. Eine multidimensionale Betrachtung. Wiesbaden 2012
- HUEBENER, Mathias; MARCUS, Jan: Empirische Befunde zu Auswirkungen der G8-Schulzeitverkürzung. In: DIW Roundup: Politik im Fokus (2015a) 57 – URL: https://www.diw.de/documents/publikationen/73/diw_01.c.497161.de/diw_roundup_57_de.pdf (Zugriff: 21.07.2020)
- HUEBENER, Mathias; MARCUS, Jan: Moving up a Gear: The Impact of Compressing Instructional Time into Fewer Years of Schooling. In: DIW-Discussion Papers (2015b) 1450 – URL:

https://www.diw.de/documents/publikationen/73/diw_01.c.496890.de/dp1450.pdf
(Zugriff: 21.07.2020)

- HURRELMANN, Klaus; SCHULTZ, Tanjev: Jungen als Bildungsverlierer: Brauchen wir eine Männerquote in Kitas und Schulen. Weinheim 2012
- KLEMM, Klaus: Expertise zur Diskussion um die Dauer der gymnasialen Schulzeit in Bayern – URL: <https://docplayer.org/9461197-Klaus-klemm-expertise-zur-diskussion-um-die-dauer-der-gymnasialen-schulzeit-in-bayern.html> (Zugriff: 09.03.2021)
- KRAMER, Jochen u. a.: Die Klasse an die Universität, die Masse an die anderen Hochschulen? Wie sich Studierende unterschiedlicher Hochschultypen unterscheiden. In: Zeitschrift für Erziehungswissenschaft 14 (2011) 3, S. 465–487 – URL: <http://dx.doi.org/10.1007/s11618-011-0213-4> (Zugriff: 21.07.2020)
- KRAUSE, Lea: Die Generation Y – ihre Wünsche und Erwartungen an die Arbeitswelt. München 2015
- KÜHN, Svenja Mareike: Sind 12 Schuljahre ausreichend für den Zugang zur Hochschule? Der doppelte Abiturjahrgang aus empirischer Perspektive. In: Beiträge zur Hochschulforschung 36 (2014) 3, S. 8–33 – URL: https://www.pedocs.de/volltexte/2015/10131/pdf/Beitraege_zur_Hochschulforschung_2014_3_Kuehn.pdf (Zugriff: 21.07.2020)
- KÜHN, Svenja Mareike: Der Streit um „G8“: Kürzere Schulzeit, mehr Stress, weniger Bildung? Positionen und Befunde zur Schulzeitdebatte – URL: <http://www.bpb.de/gesellschaft/kultur/zukunft-bildung/211182/g8-versus-g9?p=all> (Zugriff: 07.12.2016)
- KÜHN, Svenja Mareike u. a.: Wie viele Schuljahre bis zum Abitur? Eine multiperspektivische Standortbestimmung im Kontext der aktuellen Schulzeitdebatte. In: Zeitschrift für Erziehungswissenschaft 16 (2013) 1, S. 115–136 – URL: https://www.pedocs.de/volltexte/2015/10132/pdf/ZfE_2013_1_Kuehn_et_al_Wie_viele_Schuljahre_bis_zum_Abitur.pdf (Zugriff: 21.07.2020)
- KULTUSMINISTERKONFERENZ (Hrsg.): Dauer der Schulzeit bis zum Abitur. G 8 und G 9 in den Ländern – URL: <https://www.kmk.org/themen/allgemeinbildende-schulen/bildungswege-und-abschluesse/sekundarstufe-ii-gymnasiale-oberstufe-und-abitur.html> (Zugriff: 11.04.2017)
- KUPFER, Franziska: Duale Studiengänge aus Sicht der Betriebe – Praxisnahes Erfolgsmodell durch Bestenauslese. In: Berufsbildung in Wissenschaft und Praxis 47 (2013) 4, S. 25–29 – URL: <https://www.bwp-zeitschrift.de/de/bwp.php/de/bwp/show/7105> (Zugriff: 21.07.2020)
- LADENTHIN, Volker; KRAUTZ, Jochen; NOSTADT, Anja: Weniger ist weniger: G8 und die Kollateralschäden – Analysen und Materialien. Bonn 2016
- LEBRENZ, Christian; REGNET, Erika: Arbeitgeberattraktivität 2013. Betriebsklima vor Gehalt: Was macht Arbeitgeber interessant? Die Sicht der Absolventen der Hochschule Augsburg – URL: http://www.hs-augsburg.de/~eregnet/Veroeffentlichungen/Arbeitgeberattraktivitaet_2013_Broschuere_gesamt.pdf (Zugriff: 17.09.2016)

- LEVEN, Ingo; QUENZEL, Gudrun; HURRELMANN, Klaus: Familie, Bildung, Beruf, Zukunft: Am liebsten alles. In: SHELL DEUTSCHLAND HOLDING (Hrsg.): JUGEND 2015. Eine pragmatische Generation im Aufbruch. Frankfurt am Main 2015, S. 47–110
- MEYER, Tobias; THOMSEN, Stephan: Are 12 years of schooling sufficient preparation for university education? Evidence from the reform of secondary school duration in Germany – URL: http://www.niw.de/uploads/pdf/publikationen/NIW_Discussion_Paper_8_revised.pdf (Zugriff: 10.09.2016)
- MILDE-BUSCH, Astrid u. a.: Besteht ein Zusammenhang zwischen der verkürzten Gymnasialzeit und Kopfschmerzen und gesundheitlichen Belastungen bei Schülern im Jugendalter? In: Klinische Pädiatrie (2010) 222, S. 255–260 – URL: https://www.dmkg.de/files/dmkg.de/Aerzte/KlinPadiatr_0132.pdf (Zugriff: 21.07.2020)
- MINISTERIUM FÜR SCHULE UND WEITERBILDUNG DES LANDES NORDRHEIN-WESTFALEN (Hrsg.): Weiterentwicklung des Gymnasiums (G8) – URL: <https://www.schulministerium.nrw.de/docs/Schulpolitik/G8/> (Zugriff: 07.12.2016)
- MINKLEY, Nina u. a.: Mehr Stress durch G8? Stressbelastung von Abiturienten mit regulärer und verkürzter Gymnasialzeit in NRW. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 47 (2015) 4, S. 188–198 – URL: https://www.cog.psy.ruhr-uni-bochum.de/cog/mam/paper/2015/minkley_stressg8_zepp_2015_.pdf (Zugriff: 09.03.2021)
- MISCHEL, Walter: The Marshmallow Test: Mastering Self-Control. New York 2014
- MORGAN, Stephen; WINSHIP, Christopher: Counterfactuals and causal inference: Methods and principles for social research. Cambridge 2007
- MÜLLER, Christoph Emanuel: Kausale Wirkungsevaluation zwischen methodischem Anspruch und empirischer Praxis. In: STOCKMANN, Reinhard; MEYER, Wolfgang (Hrsg.): Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven. Münster 2017, S. 205–222
- MULTURS, Frank; BARGEL, Tino; RAMM, Michael: Studiensituation und studentische Orientierungen. 10. Studierendensurvey an Universitäten und Fachhochschulen. Bonn, Berlin 2008
- OSTENDORF, Fritz: Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit. Regensburg 1990
- PAN, Wei; BAI, Haiyan: Propensity score analysis: fundamentals and developments. New York 2015
- PARMENT, Anders: Die Generation Y. Mitarbeiter der Zukunft motivieren, integrieren, führen. 2. Aufl. Wiesbaden 2013
- PUSTELNIK, Kolja; HALVERSCHEID, Stefan: Längsschnittliche Vergleiche von Studierenden der Mathematik und Physik in Vorkursen und im ersten Studienjahr. In: HOPPENBROCK, Axel u. a. (Hrsg.): Mathematik im Übergang Schule/Hochschule und im ersten Studienjahr. Kassel 2013, S. 122–123

- RAMM, Michael u. a.: Studiensituation und studentische Orientierungen. 12. Studierenden-survey an Universitäten und Fachhochschulen. Langfassung. Berlin 2014
- ROSENBAUM, Paul: *Observational Studie*. 2. Aufl. New York 2002
- ROSENBAUM, Paul; RUBIN, Donald: The central role of propensity score in observational studies for causal effects. In: *Biometrika* 70 (1983) 1, S. 41–55
- RUTHUS, Julia: *Employer of Choice der Generation Y. Herausforderungen und Erfolgsfaktoren zur Steigerung der Arbeitgeberattraktivität*. Wiesbaden 2013
- SCHAVAN, Annette; AHNEN, Doris: Abitur nach 12 Schuljahren? Pro und Contra. In: *Forschung und Lehre* (2001) 8, S. 472–473
- SCHOLZ, Christian: *Generation Z: Wie sie tickt, was sie verändert und warum sie uns alle ansteckt*. Weinheim 2014
- SCHULZE HEULING, Lydia; WILD, Steffen: Re-Evaluation des D21-Digital-Indexes anhand des Birnbaum-Modells. Postervortrag an der 10. Tagung der Sektion Empirische Bildungsforschung in der DGfE (AEPF/KBBB 2019 an der WWU Münster). Münster 2019
- SHELL DEUTSCHLAND HOLDING (Hrsg.): *JUGEND 2015. Eine pragmatische Generation im Aufbruch*. Frankfurt am Main 2015
- SOMMER, Hans-Thilo: *Studie. Attraktive Arbeitgeber 2016. Top 100 Arbeitgeber 2016* – URL: <http://www.berufstart.de/karriere/attraktive-arbeitgeber/studie-attraktive-arbeitgeber-2016.pdf> (Zugriff: 11.04.2017)
- STATISTISCHES LANDESAMT BADEN-WÜRTTEMBERG (Hrsg.): *Studierende an baden-württembergischen Hochschulen im Wintersemester 2014/15* – URL: https://www.statistik-bw.de/Service/Veroeff/Statistische_Berichte/323415001.pdf (Zugriff: 09.03.2017)
- STECKL, Miriam; SIMSHÄUSER, Ulla; NIEDERBERGER, Marlene: Arbeitgeberattraktivität aus Sicht der Generation Z. Eine quantitative Befragung zur Bedeutung gesundheitsrelevanter Dimensionen im Betrieb. In: *Prävention und Gesundheitsförderung* 14 (2019) 3, S. 212–217
- THIEL, Hendrik; THOMSEN, Stephan L.; BÜTTNER, Bettina: Variation of learning intensity in late adolescence and the effect on personality traits. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 177 (2014) 4, S. 861–892
- TRAUTWEIN, Ulrich u. a.: Studieren an der Berufsakademie oder an der Universität, Fachhochschule oder Pädagogischen Hochschule? Ein Vergleich des Leistungsstands, familiären Hintergrunds, beruflicher Interessen und der Studienwahlmotive von (künftigen) Studierenden aus Baden-Württemberg. In: *Zeitschrift für Erziehungswissenschaft* (2006) 9, S. 393–412
- VOSGERAU, Klaus: *Studentische Sozialisation in Hochschule und Stadt: Theorie und Wandel des Feldes. Mit einer Fallstudie zur fachspezifischen Erfahrung der Großstadt*. Frankfurt am Main 2005
- WISSENSCHAFTSRAT (Hrsg.): *Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt. Zweiter Teil der Empfehlungen zur Qualifizierung von Fachkräften vor dem*

- Hintergrund des demographischen Wandels – URL: <https://www.wissenschaftsrat.de/download/archiv/4925-15.pdf> (Zugriff: 02.04.2017)
- WILD, Steffen; SCHULZE HEULING, Lydia: How do the digital competences of students in vocational schools differ from those of students in cooperative higher education institutions in Germany? In: Empirical Research in Vocational Education and Training (2020) 12 (5), S. 1–18 – URL: <https://doi.org/10.1186/s40461-020-00091-y>
- WOLTER, Andrä: Der Ort des dualen Studiums zwischen beruflicher und akademischer Bildung: Mythen und Realität. In: FASSHAUER, Uwe; SEVERING, Eckart (Hrsg.): Verzahnung beruflicher und akademischer Bildung. Duale Studiengänge in Theorie und Praxis. Bonn 2016, S. 39–60
- WOLTER, Andrä u. a.: Potenziale des dualen Studiums in den MINT-Fächern. Eine empirische Untersuchung (acatech STUDIE) – URL: <https://www.acatech.de/publikation/potenziale-des-dualen-studiums-in-den-mint-faechern-eine-empirische-untersuchung/> (Zugriff: 09.03.2021)
- WÖSSMANN, Ludger u. a.: Was die Deutschen über die Bildungspolitik denken – Ergebnisse des ersten ifo Bildungsbarometers – URL: <https://www.ifo.de/publikationen/2014/aufsatz-zeitschrift/was-die-deutschen-ueber-die-bildungspolitik-denken> (Zugriff: 09.03.2021)

Junmin Li

► **Ex-ante-Evaluation als Beitrag zum Policy-Transfer? Eine Analyse am Beispiel der Übertragung des Peer-Review-Verfahrens der Berufsbildung von Deutschland nach China**

Unterschiedliche Studien zeigen, dass landesspezifische Adaptionen während eines Policy-Transfers immer zu berücksichtigen sind. Welche Adaptionen tatsächlich notwendig sind, wird häufig erst während der Übertragung erfasst. Der Beitrag widmet sich der Frage, inwieweit die Methode der Ex-ante-Evaluation für den Policy-Transfer eingesetzt werden kann, um Herausforderungen noch vor dem Policy-Transfer einzuschätzen und Adaptionen zu eruieren. Die Frage wird anhand einer Ex-ante-Evaluation zu der Übertragbarkeit des Qualitätserhebungsmodells „Peer Review“ für berufsbildende Schulen beantwortet. Der Beitrag schlussfolgert, dass die Ex-ante-Evaluation Adaptionsbedarfe auf der kulturellen, berufsbildungsspezifischen sowie programmspezifischen Dimension identifizieren kann.

1 Problemstellung

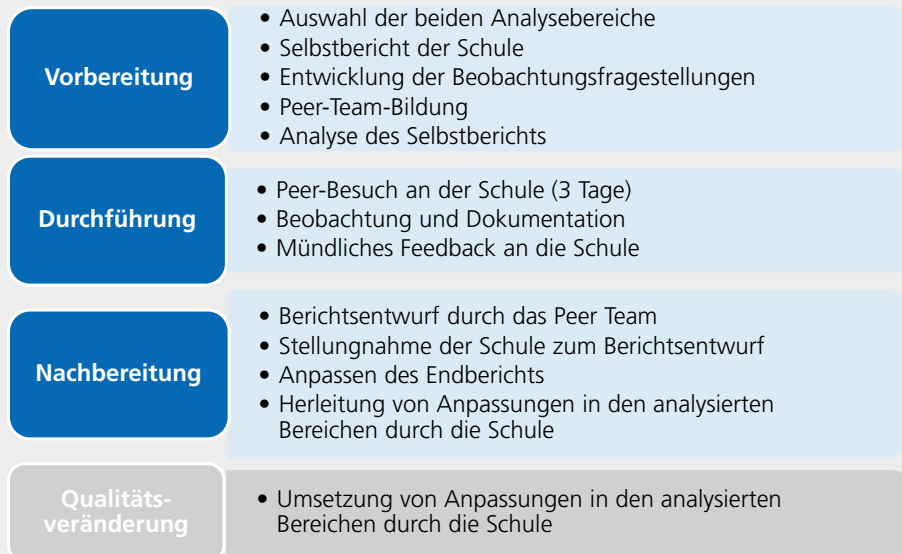
In der heutigen globalen Welt werden nicht nur Güter und Technologien über Landesgrenzen und Kontinente hinweg exportiert, sondern auch Wissen, Verfahren sowie Konzepte. Letzteres wird als Policy-Transfer bezeichnet. Unter Policy wird die inhaltliche Dimension der Politik wie etwa Reformen, Einzelmaßnahmen und Regelungen zusammengefasst (vgl. DOLOWITZ/MARSH 2000, S. 5). Dolowitz und Marsh (2000, S. 5) definieren Policy-Transfer als die Übertragung dieser Policies von einem politischen in ein anderes politisches System. Im deutschen Kontext findet Policy-Transfer in der Berufsbildung im Rahmen von Entwicklungszusammenarbeit sowie internationaler Zusammenarbeit statt (vgl. BMBF 2016; FRAUNHOFER MOEZ 2012). Unterschiedliche Studien zeigen, dass die Eins-zu-eins-Übertragung von fremden Verfahren und Konzepten von einem Land in ein anderes selten möglich ist. Stattdessen sind landesspezifische Adaptionen des Transferobjektes im Policy-Nehmer-

Land notwendig (z. B. vgl. SCHIPPERS 2009; SILVESTRINI/STOCKMANN 2019). Wissenschaftliche Evaluationen, wie sie von Schippers (2009), Silvestrini und Stockmann (2013) sowie Stockmann und Meyer (2017) durchgeführt worden sind, untermauern solche Adaptionsbedarfe mit empirischen Daten. Erkenntnisse über konkrete Adaptionsnotwendigkeit werden häufig erst nach einem abgeschlossenen Transferprojekt durch die Konfrontation mit konkreten Herausforderungen erfasst. Folgen solcher Konfrontationen sind ungeplanter, erhöhter Ressourcenverbrauch (z. B. Projektgelder und Personaleinsatz im Rahmen der Umsetzung) und Verbesserungsschleifen, die sogar zum Scheitern eines Policy-Transfers führen. Die Ex-ante-Evaluationsforschung nimmt an, dass sich ein erhöhter Aufwand reduzieren lässt, wenn Herausforderungen bereits vor der Implementation identifiziert werden können (vgl. SILVESTRINI 2011; MOSSBERGER/WOLMAN 2003). Der Beitrag beantwortet die Frage, bis zu welchem Ausmaß die Methode der Ex-ante-Evaluation im Rahmen des Policy-Transfers eingesetzt werden kann, um die notwendigen landesspezifischen Adaptionen vor dem Policy-Transfer zu eruieren und somit eine systematische und elaborierte Übertragung zu ermöglichen. Konkret wird die Fragestellung anhand einer Ex-ante-Evaluation zur der Übertragbarkeit des Qualitätserhebungsmodells „Peer-Review“ für berufsbildende Schulen von Deutschland nach China beantwortet. Die Forschungsergebnisse zeigen, dass die Ex-ante-Evaluation konkrete Einschätzungen zu fördernden und hemmenden Faktoren während des Policy-Transfers geben kann.

2 Peer-Review-Verfahren für berufsbildende Schulen in China

Das Peer-Review-Verfahren für berufsbildende Schulen ist ein Instrument der Qualitätssicherung und -entwicklung. Das Verfahren wurde durch diverse Modellversuche auf der europäischen Ebene erprobt. Auch in Deutschland gab es Modellversuche zum Einsatz von Peer-Review-Verfahren in berufsbildenden Schulen (vgl. BUHREN/GIESKE/ROLFF 2012). Das Peer-Review-Verfahren spiegelt westliche demokratische Werte der Partizipation wider (vgl. GUTKNECHT-GMEINER 2008), denn das Verfahren unterstützt die Schule in ihrer eigenen Bemühung um eine maßgeschneiderte Schulentwicklung. Ähnlich wie bei dem Peer-Review-Verfahren im Hochschulwesen agieren hier Peers, d. h. externe, gleichgestellte und unabhängige Kollegen und Kolleginnen als Beobachtende, die den Schulen ihre Wahrnehmungen zu spezifischen Fragestellungen mitteilen. Das Verfahren verfolgt einen partizipatorischen Ansatz und analysiert den Zustand der von der Schule selbst bestimmten Analysebereiche. Dabei ist das Verfahren in die drei Phasen Vorbereitung, Durchführung sowie Nachbereitung untergliedert (siehe Abbildung 1).

Abbildung 1: Die drei Phasen des Peer-Review-Verfahrens und die daran anschließende Qualitätsveränderung



Quelle: eigene Darstellung

Als Vorbereitung auf den Peer-Besuch wählt die Schule zwei Analysebereiche für das Peer-Review aus und verfasst dazu einen Selbstbericht. Zusätzlich schlägt die Schule Beobachtungsfragestellungen vor, die sie als relevant erachtet. Basierend auf den ausgewählten Analysebereichen wird das Peer-Team gebildet, welches den Selbstbericht analysiert und den Besuch vorbereitet. Die Analyse erfolgt durch einen dreitägigen Peer-Besuch. Währenddessen hospitieren die Peers im Unterricht, führen Gespräche mit Lehrkräften und Schülern bzw. Schülerinnen und analysieren Dokumente. Die Ergebnispräsentation des Peer-Besuchs findet anschließend in Form einer Wiedergabe von Beobachtungen ohne Beurteilung statt. Die Bewertung der Ergebnisse sowie die Ableitung von Konsequenzen für die Schulentwicklung liegen im Verantwortungsbereich der Schule. Somit erhält die Schule die Freiheit, das Ergebnis des Peer-Review-Verfahrens nach den schulischen Rahmenbedingungen zu bewerten und schulspezifische Maßnahmen abzuleiten (vgl. GUTKNECHT-GMEINER 2008). Empirische Studien zeigen, dass der partizipatorische Ansatz des Verfahrens zu einer hohen Akzeptanz der Ergebnisse auf allen Seiten führt. Es fördert das Verantwortungsgefühl der Lehrkräfte für die schulische Qualitätsentwicklung. Allerdings setzt dieses Verfahren eine hohe intrinsische Motivation der schulischen Akteure voraus, welche durch strukturelle und individuelle Faktoren beeinflusst wird. Die Literatur zeigt ebenfalls, dass die Schule für eine

tatsächliche Qualitätssteigerung zeitliche, personelle und finanzielle Ressourcen benötigt (vgl. GUTKNECHT-GMEINER 2008; WULZ/JONACH/GRAMLINER 2011).

Die Besonderheiten des Peer-Review-Verfahrens könnten den derzeit in China vorherrschenden Herausforderungen der Qualitätsentwicklung in der Berufsbildung entgegenwirken. In der Kritik stehen vor allem hierarchische, starre und bürokratische Evaluationsverfahren durch hochrangige Autoritätspersonen der Bildungsministerien. Die berufsbildenden Schulen Chinas verfügen über wenig Mitgestaltungsrechte, und es werden Bewertungsmaßstäbe verwendet sowie Handlungsempfehlungen ausgegeben, die nicht die Realität der Schulen einbeziehen (vgl. ZHANG/ZHANG 2009; YANG 2015). Falls sich das Peer-Review-Verfahren mit gewissen Anpassungen auf China übertragen ließe, könnte es Potenziale für eine offene und gleichberechtigte Evaluationskultur bieten und zu einer realistischen Qualitätsentwicklung führen.

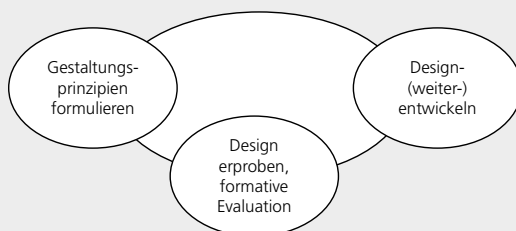
3 Ex-ante-Evaluation zur Identifizierung von Anpassungsbedarfen des Policy-Transfers

Vor dem Beginn eines Policy-Transfers stehen die Akteure oft vor der Herausforderung, dass noch keine Erfahrungswerte darüber existieren, wie die zu übertragende Policy im fremden Kontext funktioniert und wie diese mit den bereits bestehenden Strukturen interagieren wird. Es werden teilweise Spekulationen getroffen, welche intendierten und nicht intendierten Wirkungen auftreten könnten. Eine Ex-ante-Evaluation kann zu wissenschaftlich fundierten Einschätzungen von Wirkungen und Nebenwirkungen verhelfen und willkürliche Spekulationen eliminieren (vgl. ROSE 2005). Denn diese spezifische Form der Evaluation, auch als prospektive Evaluation bezeichnet, hat das Ziel, die Wirkung einer Policy nach wissenschaftlichen Standards einzuschätzen, noch bevor diese eingeführt wird. Sie setzt die systematische Sammlung und Auswertung von organisationalen und prozessualen Daten ein, um die Wirkung einer Policy vorherzusagen (vgl. MOSSBERGER/WOLMAN 2003; SILVESTRINI 2011). Dafür werden Rahmenbedingungen analysiert und Befunde aus ähnlichen vergangenen Studien einbezogen. Die Ex-ante-Evaluation dient somit als eine Art „Früherkennungssystem“, als „an early warning system about the risk of failure“ (ROSE 2005, S. 132), und wirkt auf die Gestaltung des Policy-Transfers ein.

Die Methode der Ex-ante-Evaluation beinhaltet zum einen die Interpretation von historischen Daten, zum anderen die Generierung von Primärdaten durch Experimente, Simulationen, Meinungsumfragen, Befragungen von Fachleuten sowie Szenarienbildung und Pilotstudien (vgl. WOLPIN 2007; MOSSBERGER u. a. 2003, SILVESTRINI 2011; LEWIS 2007). Da im Policy-Nehmerland China kaum Erfahrungswerte zu partizipativer Qualitätsentwicklung vorhanden sind, um eine fundierte Einschätzung abzugeben, wird die Ex-ante-Evaluation im Rahmen einer Pilotstudie umgesetzt. Die Besonderheit dieser Ex-ante-Evaluation ist der Einsatz des iterativen Mikrozyklus des *Design-Based-Research*. Der *Design-Based-Research* ist ein gestaltungsorientierter Ansatz aus der Lehr-/Lernforschung zur Erprobung von Innovatio-

nen. Durch eine mehrzyklische Erprobung eines didaktischen Konzeptes wird das Konzept systematisch angepasst. Die mehrzyklische Erprobung wird durch den iterativen Mikrozyklus realisiert (vgl. DILGER 2014). Der iterative Mikrozyklus besteht aus drei Phasen. Die erste Phase des iterativen Mikrozyklus beginnt mit der Entwicklung eines Konzeptes. Dann wird dieses Konzept in der zweiten Phase erprobt und formativ evaluiert. Abschließend werden in der dritten Phase auf Basis der formativen Evaluation Gestaltungsprinzipien entwickelt, die wiederum in die erste Phase der Konzeptentwicklung einfließen (siehe Abbildung 2) (vgl. DILGER 2014).

Abbildung 2: Iterativer Mikrozyklus des *Design-Based-Research*



Quelle: eigene Darstellung in Anlehnung an DILGER 2014, S. 366

Die Anwendung des iterativen Mikrozyklus ermöglicht eine systematische Erhebung von Anpassungsbedarfen sowie Anpassungserprobungen bereits vor der eigentlichen Implementation. Herausforderungen, die während der konkreten Pilotstudie identifiziert werden, dienen als Referenz für eine Anpassung des Peer-Review-Verfahrens. Zudem verfolgt der iterative Mikrozyklus einen partizipativen Gestaltungsansatz. Die beteiligten Lehrkräfte, Schulleitungspersonen sowie die Peers werden in die Ableitung von Anpassungen einbezogen, sodass die aus der Ex-ante-Evaluation geschlussfolgerten Anpassungsvorschläge auf Akzeptanz der einheimischen Akteure stoßen.

Die Ex-ante-Evaluation baut auf der Programmtheorie des Peer-Review-Verfahrens auf, welche eigens für dieses Vorhaben entwickelt wurde. Die Programmtheorie dokumentiert die Kausalität des Peer-Review-Verfahrens (vgl. BALZER 2005). Im ersten Schritt sind die mittelfristigen Ziele anhand der Literatur zum Peer-Review-Verfahren zu identifizieren, bevor im zweiten Schritt Outcomes, die die Zielerreichung indizieren, festgehalten werden. Outcomes sind Indikatoren, die anzeigen, inwieweit ein Ziel erreicht wird. Dabei kann ein Ziel über mehrere Outcomes verfügen. In den anschließenden Schritten werden entsprechend der Programmevaluation die Determinanten und Interventionen festgelegt (vgl. CHEN 2005).

Im vorliegenden Fall wurden insgesamt neun Outcomes, 15 Determinanten und 32 Interventionen identifiziert. Tabelle 1 stellt die Beziehung der mittelfristigen Ziele und Outcomes des Peer-Review-Verfahrens zusammenfassend dar. Hier wird deutlich, dass ein Outcome teilweise als Indikator für die Erreichung mehrerer Ziele fungiert.

Tabelle 1: Mittelfristige Ziele und Outcomes des Peer-Review-Verfahrens

Ziele	Outcomes
Betroffene und Beteiligte werden zu Akteuren der schulischen Qualitätssicherung.	Fokus auf die am Verfahren beteiligten Menschen, deren Interessen und Bedürfnisse setzen
	Offenheit und Ehrlichkeit zwischen allen Beteiligten schaffen
	eine hinterfragende und kritische Haltung sowohl bei den Peers als auch bei den Beteiligten der Schule fördern
Der Rechtfertigungsdruck der Schule gegenüber Qualitätserhebungsaktivitäten wird gesenkt.	Vertraulichkeitsvereinbarungen und Regeln über die Verwendung der Ergebnisse einhalten
	Interessenkonflikte und direkte Konkurrenz zwischen den Peers (bzw. deren Herkunftsschule) und der Schule vermeiden
	Offenheit und Ehrlichkeit zwischen allen Beteiligten schaffen
	das Peer-Review-Verfahren dynamisch und flexibel gestalten
Die Möglichkeit des Voneinander-Lernens zwischen allen Beteiligten wird geschaffen.	Transparenz in allen Bereichen des Verfahrens für alle beteiligten Personen herstellen
	Offenheit und Ehrlichkeit zwischen allen Beteiligten schaffen
	eine hinterfragende und kritischen Haltung sowohl bei den Peers als auch bei den Beteiligten der Schule fördern
Eine objektive externe Sicht auf die Analysebereiche wird ermöglicht.	Objektivität und Unparteilichkeit der Peers gewährleisten
	Transparenz in allen Bereichen des Verfahrens für alle beteiligten Personen herstellen
	Interessenkonflikte und direkte Konkurrenz zwischen den Peers (bzw. deren Herkunftsschule) und der Schule vermeiden
	Offenheit und Ehrlichkeit zwischen allen Beteiligten schaffen
	Feedback für die Schulentwicklung nutzen

Um einen Outcome zu erreichen, werden Determinanten und Interventionen definiert. Determinanten sind Zustände, die erreicht werden müssen, damit ein Outcome erreicht werden kann. Die Intervention ist eine Aktivität, die durchzuführen ist, damit ein Zustand bzw. eine Determinante entsteht. Eine Determinante für den Outcome „Offenheit und Ehrlichkeit zwischen allen Beteiligten schaffen“ ist „offene, gleichgestellte und rege Kommunikation zwischen dem Peer-Team und den Beteiligten der evaluierten Schule“. Die Intervention „das Peer-Team gibt kein eigenes Urteil und keine Handlungsempfehlung ab, sondern gibt der evaluierten Einrichtung ein deskriptives Feedback der Beobachtung“ führt zu dieser Determinante.

4 Durchführung der Ex-ante-Evaluation

Die Pilotstudie im Rahmen der Ex-ante-Evaluation wurde in vier berufsbildenden Schulen in Shanghai im Jahr 2014 realisiert. Die Schulen sind in den Bereichen Verwaltung, Pflege, Kunsthandwerk und Gewerbe zu verorten. In der Vorbereitungsphase bestimmten die Schulen jeweils zwei Analysebereiche, die vom Peer-Team untersucht werden sollten (siehe Tabelle 2).

Tabelle 2: Ausgewählte Analysebereiche der Schulen für das Peer-Review-Verfahren

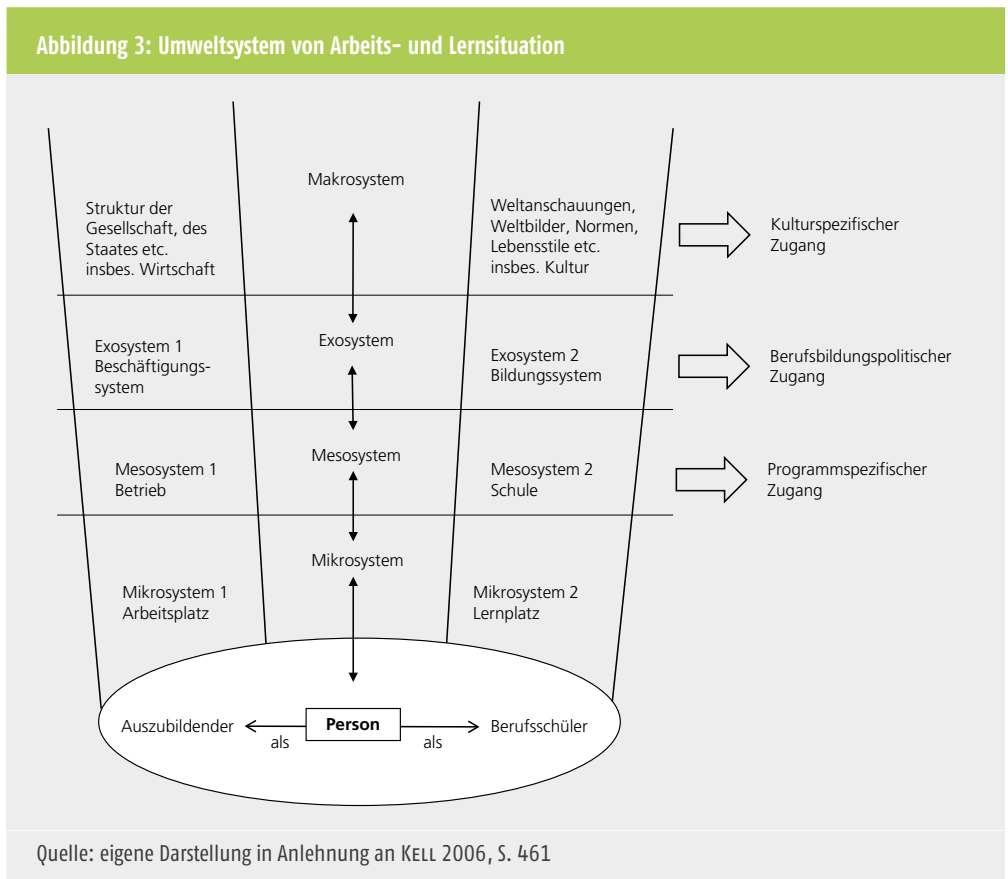
Schule	Analysebereiche
Schule 1	Innovationsprojekt im Bereich Übergang Schule – Beruf Berufsorientierungsunterricht
Schule 2	Unterricht im Fach Schmuckdesign Unterricht zur Vorbereitung von Bewerbungsunterlagen
Schule 3	Unterricht im Bereich der medizinischen Grundkenntnisse Englisch für Krankenpfleger/-innen
Schule 4	Unterricht zum Thema Kabellegen Praxisunterricht im Fach Instandhaltung von Elektronik

Für die Ex-ante-Evaluation wurden die Methoden der teilnehmenden Beobachtung sowie des problemzentrierten Interviews verwendet. Die teilnehmende Beobachtung wurde genutzt, um den Prozess des Transfers zu analysieren und zu dokumentieren (vgl. LAMNEK 1993). Die Kernbeobachtungszeit für die Durchführung der vier Peer-Besuche betrug fünf Wochen. Die Beobachtungen wurden mithilfe von Beobachtungsprotokollen festgehalten. Hinzu kommt die Dokumentation der Vorbereitungs- sowie der Nachbereitungszeiten, die nur schwer zu quantifizieren sind. Die Methode des qualitativen problemzentrierten Interviews diente zur Erhebung der Akzeptanz, Wahrnehmung und Einstellungen zum Peer-Review-Verfahren (vgl. WITZEL 2000). Es wurden Interviews mit acht fachlich versierten Peers, acht Lehrkräften der evaluierten Schulen, die als Koordinatoren bzw. Koordinatorinnen agierten und vier Schulleitungspersonen geführt. Ein Interview dauerte im Durchschnitt 60 Minuten und wurde digital aufgezeichnet und transkribiert. Beide Methoden waren als komplementär zu verstehen: Die teilnehmende Beobachtung evaluierte den Prozess der Pilotstudie und das problemzentrierte Interview evaluierte die Zielerreichung. Da auch während der Interviews Fragen zum Verlauf des Peer-Review-Verfahrens gestellt wurden, gab es eine Schnittmenge, die durch beide Methoden untersucht wurde. Hier nutzte die Studie die wissenschaftliche Triangulation.

Sowohl für die Beobachtungsdimensionen als auch für den Interviewleitfaden diente die Programmtheorie des Peer-Review-Verfahrens als Grundlage, sodass eine Rückkopplung der Untersuchungsergebnisse mit den Programmzielen des Verfahrens gewährleistet werden konnte. Die Auswertung fand mit der qualitativen Inhaltsanalyse statt (vgl. MAYRING 2003).

5 Interpretationsrahmen für die Evaluationsergebnisse

Die Ergebnisse der Ex-ante-Evaluation wurden in Anlehnung an das Modell der Umweltsysteme von Arbeits- und Lernsituation nach Kell (2006, S. 461) interpretiert (siehe Abbildung 3).



Das Modell nach Kell (2006) unterscheidet zwischen den Lernorten Ausbildungsbetrieb und berufsbildende Schule. Hier wurden aufgrund des Fokus des Peer-Review-Verfahrens auf die schulische Qualitätsentwicklung nur die schulischen Umweltebenen betrachtet. Gemäß dem Modell nutzte das Forschungsvorhaben auf der Makroebene den kulturellen Zugang für die

Ergebnisinterpretation. Das hier verwendete Kulturverständnis lehnt sich an das Orientierungssystem nach Herrmann-Pillath (1997) an. Herrmann-Pillath (1997, S. 2) gliedert das chinesische Kulturkonzept in zwei wesentliche Bestandteile: Er unterscheidet zwischen Werten und Weltsicht einerseits sowie Kommunikation und Interaktion andererseits. Die Werte und Weltsicht beeinflussen die Kommunikation und Interaktion. Die Werte und Weltsicht der chinesischen Kultur werden hier anhand der drei traditionellen Lehren Chinas Konfuzianismus, Taoismus und Buddhismus erörtert. Der Einfluss dieser Lehren prägt die Traditionen und Praktiken Chinas und ist bis in die moderne Zeit sichtbar (vgl. YU/TAYLOR/TUNG 2003, S. 171; CLART 2009, S. 162). Es sind darüber hinaus weitere Einflüsse zu nennen, wie etwa die politische Geschichte Chinas (vgl. HEILMANN 2002), die hier berücksichtigt werden. Auf Basis der genannten Einflussgrößen wurden Kommunikations- und Interaktionslinien abgeleitet, die teilweise in den Kulturkonzepten von Hall (1959; 1966; 1976; HALL/HALL 1990) sowie Hofstede/Hofstede/Minkov (2010) wiederzufinden sind. Die Kulturkonzepte der genannten Autoren und Autorinnen werden hier jedoch kritisch reflektiert. Die resultierenden Kommunikations- und Interaktionslinien lassen sich u. a. in die Dimensionen *In-Group*- und *Out-Group*-Orientierung, hohe Machthierarchie, Maskulinität, hoher Kontextbezug der Kommunikation, polichrones Zeitmanagement, Langfristorientierung, Selbstbeherrschung sowie passive Haltung gegenüber Veränderung untergliedern.

Auf Ebene des Exosystems wurde der berufsbildungspolitische Zugang gewählt. Die Entwicklung des chinesischen Berufsbildungssystems sowie dessen Aufbau wurden hier beleuchtet. Zudem dienen die gegenwärtigen Herausforderungen der chinesischen Berufsbildung als Interpretationszugang für die Ergebnisse. Konkret wurden die regional unterschiedliche Entwicklung zwischen West- und Ostchina (vgl. z. B. LIU/JIANG/TROWE 2006), der gesellschaftliche Status der Berufsbildung (vgl. z. B. VELDE 2009) sowie der Qualitätsmangel der Berufsbildung (vgl. z. B. HAO 2012) für die Interpretation einbezogen.

Das Qualitätsmanagementkonzept ist als programmspezifischer Interpretationsrahmen dem Mesosystem des Modells nach Kell (2006) zuzuordnen. Hier wurde die Entwicklung der Qualitätssicherung in der Berufsbildung untersucht (vgl. z. B. YAO 2008; ZHANG 2013; MoE 2003). Die konkrete Erschließung der Untersuchungsergebnisse basierte auf dem Qualitätsmanagementkonzept der berufsbildenden Schulen in Shanghai (vgl. SHANGHAI MUNICIPAL EDUCATION COMMISSION 2009). Des Weiteren wurde das Peer-Review-Verfahren als Programm auf dieser Ebene analysiert. Die Ergebnisse der Pilotstudie wurden mit den Studienergebnissen zum Peer-Review-Verfahren im europäischen und deutschen Kontext gegenübergestellt, um programmspezifische Erklärungen zu identifizieren, die unabhängig von landesspezifischen Rahmenbedingungen waren (vgl. BUICHL/WILBERS 2011; GUTKNECHT-GMEINER 2010; BUHREN/GIESKE/ROLFF 2012).

Das Mikrosystem des Lernplatzes wurde für die Interpretation der Ergebnisse ausgeschlossen, da der Fokus des Peer-Review-Verfahrens nicht auf den einzelnen Unterrichtsstunden liegt. Die Ergebnisse der Ex-ante-Evaluation zeigen, dass die kulturellen Faktoren die wichtigsten Erklärungen für die Interpretation der Ergebnisse liefern.

5.1 Kulturspezifische Interpretation

Eine Besonderheit des Peer-Review-Verfahrens ist die partizipative Zusammenarbeit zwischen dem Peer-Team und der Schule. Auf kultureller Ebene erschwerte die starke hierarchische Machtstruktur der Schule die partizipative Zusammenarbeit. Beispielsweise weigerten sich die Schulleitungen aller vier Schulen, das Kollegium in die Auswahl der zu untersuchenden Analysebereiche einzubeziehen und bestimmten diese Bereiche selbst. Eine Lehrkraft antwortete während des Interviews auf die Frage, weshalb ihr Unterrichtsbereich ausgewählt wurde, wie folgt: „Ganz genau, weshalb diese Analysebereiche ausgewählt wurden, weiß ich nicht. Es war die Entscheidung der Abteilung für Lehre“ (Lehrkraft 1).

Auch zeigen die Beobachtungen, dass die Lehrkräfte der Schulen sich in der Zusammenarbeit mit dem Peer-Team ausschließlich reaktiv verhielten. Trotz der mehrmaligen Aufforderungen des Peer-Teams brachten die Lehrkräfte keine eigenen Ideen oder Vorschläge für die Gestaltung der Peer-Besuche ein. Es konnte dennoch in der Zusammenarbeit, bis auf eine Ausnahme, ein kollegialer, wohlwollender und gleichgestellter Umgang zwischen den Peers und den Lehrkräften der Schulen hergestellt werden. Eine Lehrkraft beschrieb ihren Eindruck wie folgt:

„Wir hatten das Gefühl, dass die fachlich versierten Peers in deren Schule teilweise auf dieselben Herausforderungen wie wir gestoßen sind. Deswegen empfanden wir das Gespräch als anregend. [...] Wir konnten über sehr vieles reden. Deswegen fand ich, dass die Diskussion mit den Peers an diesem Tag sehr erfolgreich war. Alle waren ehrlich, es wurde nichts vorenthalten.“ (Lehrkraft 4)

Bei der erwähnten Ausnahme agierten ehemalige, bereits pensionierte Schulleiter anstatt Lehrkräfte einer berufsbildenden Schule als Peers. Diese Personen wurden von der evaluierten Schule mit besonderer Wertschätzung und autoritätsüblichen Gepflogenheiten behandelt, die bei den anderen drei Schulen nicht zu beobachten waren. Trotz mehrmaliger Betonung des Peer-Teams, eine kollegiale und gleichgestellte Behandlung zu bevorzugen, wurden sie dennoch weiterhin gesondert behandelt. Hier zeigt sich die Wichtigkeit der informellen Hierarchieordnung sowie des Senioritätsprinzips in der chinesischen Interaktion (vgl. HOFSTEDE/HOFSTEDE/MINKOV 2010, S. 239; SHENKAR/RONEN 1990).

Die nicht wertende Rückmeldung der Beobachteten verlief aufgrund des hohen Kontextbezugs in dem chinesischen Kommunikationsstil zum größten Teil reibungslos. Die Interviews zeigen, dass die Lehrkräfte das Feedback des Peer-Teams in den Gesamtkontext der Schule einordnen und daraus Schlussfolgerungen für die Qualitätsfragen ziehen konnten.

Zugleich deutet die Ex-ante-Evaluation darauf hin, dass es aufgrund des hohen Kontextbezugs des Kommunikationsstils einer Anpassung der Kommunikationskanäle im Peer-Review-Verfahren bedarf. Die üblichen Kommunikationskanäle in der Vorbereitung des Peer-Besuchs beruhten auf einer ausschließlich verbalen, kontextarmen Informationscodierung (vgl. GÜNTNER 1993, S. 170). Diese wurde hier durch E-Mails und Telefonate erschwert.

Insbesondere kritische Informationen werden in China indirekt über den Einbezug des Kontexts, z. B. durch Gestikulation, Gesprächspausen oder Metaphern, weitergegeben, um das Gesicht zu wahren (vgl. SHI 2003, S. 97). In zwei Fällen kommunizierten die Schulleitungspersonen ihre Befangenheit gegenüber den Herkunftsschulen der fachlich versierten Peers erst in den Interviews, weil in der Vorbereitungsphase keine Möglichkeit der kontextbezogenen Kommunikation gegeben wurde. Eine Schulleitungsperson erzählte während des Interviews von ihren Bedenken:

„Wegen Konkurrenz könnte es passieren, dass die fachlich versierten Peers das Feedbackergebnis absichtlich negativ verzerren. Damit die Schule schlecht dasteht. [...] Aber die beiden fachlich versierten Peers, die bei uns waren, haben wissenschaftlich gearbeitet, weil diese fachlich versierten Peers keine leitenden Positionen innehatten. Bei Leitungen würden sie die Beziehung zwischen beiden Schulen beim Peer-Review berücksichtigen.“
(Schulleitungsperson 3)

Die Ergebnisse der Pilotstudie zeigen folglich, dass das chinesische Gesichtskonzept je nach Intervention sowohl fördernd als auch hemmend wirken kann. An dieser Stelle wird die Komplexität der chinesischen Kommunikationsform deutlich.

Die Ex-ante-Evaluation zeigt, dass mehrheitlich ein offener und ehrlicher Austausch zwischen dem Peer-Team und den Lehrkräften erreicht wurde. Die gleichgestellte und zugleich externe Eigenschaft der Peers förderte besonders die Offenheit sowie die Ehrlichkeit der Diskussionsrunden. Die Lehrkräfte hatten keinen Druck, sozialkonform antworten zu müssen. Hier wird die *In-Group*- und *Out-Group*-Orientierung sowie die Hierarchieorientierung in der chinesischen Kultur deutlich (vgl. HERRMANN-PILLATH 1997; TRIANDIS 1995; CHEN/STAROSTA 2005). Die Lehrkräfte der Schulen gewichteten die Rückmeldungen der externen Peers (*Out-Group*) höher als diejenigen von ihrer eigenen Schule (*In-Group*). Die Schüler/-innen der Schulen reagierten ebenso offener und ehrlicher gegenüber externen Lehrkräften. Auch die nicht bewertende Eigenschaft des Verfahrens führte zu einer offenen und ehrlichen Kommunikation. Beispielsweise erzählte eine Lehrkraft bezüglich des Feedbacks:

„Manche Aussagen der Schüler/-innen hatten uns überrascht. Wir reden zwar auch mit denen, aber manche Sachen sagen sie uns Lehrkräften nicht direkt. Es ist denen manchmal unangenehm. Aber die Schüler/-innen hatten gegenüber dem Peer-Team detaillierte Auskunft gegeben.“ (Lehrkraft 5)

Bereits zum Anfang der Pilotstudie war es auffällig, dass die lineare Zeitplanung des Peer-Review-Verfahrens sich in China nicht eins zu eins umsetzen ließ. In der chinesischen Kultur wird die Zeit polychron genutzt. Flexibilität und situationsorientierte Planung stehen im Vordergrund. Arbeitsschritte werden im Durchschnitt nicht früher als zwei Wochen vor der Realisierung festgelegt. Die chinesischen Beteiligten gehen mit der verfügbaren Zeit, dem Arbeitsaufwand sowie mit einer kurzfristigen Veränderung in der Planung flexibel um (vgl.

HALL/HALL 1990, S. 13–17; BOSCH/REICHENBACH/SCHMIDT 2003, S. 189). Die Beobachtungen und Interviews bestätigten das polychrone Zeitmanagement. So wurden auch die einzelnen Programmschritte des Peer-Review-Verfahrens frühestens zwei Wochen vor der Realisierung festgelegt. Die Beteiligten empfanden diese Kurzfristigkeit nicht als unangenehm, sondern sahen sie als selbstverständlich an.

Zusammenfassend lässt sich durch die Ex-ante-Evaluation feststellen, dass sich Unterschiede in der Kultur bei der Policy-Übertragung je nach Interventionen und gegebenen Situationen auf der einen Seite hemmend und auf der anderen Seite fördernd auf den Policy-Transfer auswirken. In dieser Pilotstudie werden die Kulturdimensionen „hohe Machtdistanz“, „In-Group vs. Out-Group“, „Kommunikation mit hohem Kontextbezug“ sowie „polychrones Zeitmanagement“ als wichtige Faktoren attestiert.

5.2 Berufsbildungspolitische Interpretation

Das chinesische Berufsbildungssystem und die bildungspolitische Ausrichtung führten ebenso zu Anpassungen bei der Übertragung des Peer-Review-Verfahrens. Anders als in Deutschland ist der größte Teil der chinesischen Berufsbildung vollzeitschulisch organisiert. Aufgrund der vollzeitschulischen Ausbildungsform kommen die berufsbildenden Schulen selten mit den Betrieben in den Austausch. Eine Erweiterung des Peer-Teams um Vertretende der Betriebe wurde von mehreren Lehrkräften während der Interviews als Verbesserungsmöglichkeit für das Peer-Review-Verfahren angemerkt: „Es könnte noch jemand vom Unternehmen als Peer agieren, denn die aus den Unternehmen kennen sich mit den Anforderungen der Branche besser aus“ (Lehrkraft 3).

Weitere Herausforderungen bezüglich des Berufsbildungssystems sind zumindest in dem vorliegenden Forschungsvorhaben nicht attestiert worden. Dies lässt sich dadurch erklären, dass das Peer-Review-Verfahren ein Instrument ist, welches auch im europäischen Kontext vorwiegend am Lernort Schule eingesetzt wird und folglich gut in das vollzeitschulische System der chinesischen Berufsbildung passt.

5.3 Programmspezifische Interpretation

Die Interviews zeigen, dass sowohl die Peers als auch die Beteiligten aus den Schulen das Peer-Review-Verfahren als eine gute Ergänzung zum chinesischen Qualitätsmanagementsystem der Berufsbildung ansehen, weil beide Verfahren unterschiedliche Funktionen verfolgen. Der chinesische Ansatz verfolgt nach der Meinung der Interviewten eine Qualitätskontrollfunktion. Das Peer-Review-Verfahren zielt auf die Qualitätsentwicklung ab.

Der Vergleich zwischen den Ergebnissen der chinesischen Ex-ante-Evaluation und den Studien aus dem europäischen Kontext verdeutlicht, dass einige Interventionen ähnliche fördernde und hemmende Faktoren trotz der unterschiedlichen nationalen Rahmenbedingungen aufweisen. Diese ähnlichen Ausprägungen lassen programmspezifische Ursachen vermuten (vgl. GEORG 2005). Auf der programmspezifischen Ebene wurde der Mangel an

Erfahrung der Peers mit Beobachtungsmethoden identifiziert. Die Peers berichteten während der Interviews, dass sie sich nicht sicher waren, ob sie über ausreichende Kompetenz für die Durchführung des Peer-Besuchs verfügten. Dieses Phänomen wurde in den Studien zum Peer-Review-Verfahren in Europa, z. B. in Deutschland, ebenso thematisiert und scheint unabhängig von den unterschiedlichen Kontextfaktoren der Länder zu sein.

Ein fachlich versierter Peer betrachtete eine Peer-Schulung als sinnvoll:

„Denn wenn das Verfahren sich in der Zukunft verbreiten sollte, dann müssen die Peers wissen, wie man evaluiert und was man beobachtet. Methodisch ist es notwendig. Man braucht eine vorherige Vorbereitung. Sonst könnte passieren, dass manche im Peer-Besuch vielleicht noch nicht ganz verstanden haben, worum es hier geht, was wir rückmelden müssen.“ (fachlich versierter Peer 1)

Ähnlich ist auch das Ergebnis zu bewerten, dass die Lehrkräfte zum Teil unverbindliche Handlungsempfehlungen vom Peer-Team wünschten. Auch dieses Bedürfnis wurde nicht nur in China, sondern auch in anderen Studien attestiert (vgl. Li 2019).

5.4 Einschätzung der Übertragbarkeit des Peer-Review-Verfahrens nach China

Insgesamt konnte die Mehrheit der Interventionen aus der Programmtheorie des Peer-Review-Verfahrens umgesetzt werden. Eine Minderheit der Interventionen, die sich auf Partizipation bezieht, wurde nur zum Teil realisiert. Es gab keine Intervention, die gar nicht durchgeführt werden konnte. Zusammenfassend lässt sich sagen, dass das Peer-Review-Verfahren einen positiven Einfluss auf die Qualitätsentwicklung in den Schulen und auf die Peers hatte. Die Ergebnisse zeigen, dass das Verfahren die Ängste der Lehrkräfte vor der Teilnahme an der Qualitätsbewertung verringert und ihre Bereitschaft erhöht, Verantwortung für die Qualitätsentwicklung in ihrer Schule zu übernehmen. Die Ziele, „den Rechtfertigungsdruck der Schule gegenüber Qualitätserhebungsaktivitäten zu senken“, „die Möglichkeit des Voneinander-Lernens zwischen allen Beteiligten zu schaffen“ sowie „eine objektive externe Sicht auf die Analysebereiche zu erhalten“ (siehe Tabelle 1) konnten zu einem großen Grad mithilfe von Anpassungen erreicht werden. Allein das Ziel „Betroffene und Beteiligte werden zu Akteuren der schulischen Qualitätssicherung“ war aufgrund der vorhandenen Hierarchiestruktur schwer zu erreichen.

Darüber hinaus ist noch die Begrenzung finanzieller, personeller und zeitlicher Ressourcen der Pilotstudie zu benennen, die sich teilweise hemmend auf die Zielerreichung auswirkte. So war es beispielsweise aufgrund der finanziellen Beschränkung nicht möglich, Peers von außerhalb der Stadt Shanghai einzuladen. Dies führte zur Befangenheit zweier Schulen gegenüber den Herkunftsschulen der fachlich versierten Peers, da die beiden Schulen in einem Wettbewerbsverhältnis stehen.

6 Ex-ante-Evaluation führt zur systematischen Vorbereitung von Policy-Transfer

Der Beitrag schlussfolgert, dass die Ex-ante-Evaluation sowohl hemmende als auch fördernde Faktoren für den Policy-Transfer auf verschiedenen Ebenen generieren kann. Durch den partizipatorischen Ansatz der Ex-ante-Evaluation konnten zu jeder einzelnen Intervention des Peer-Review-Verfahrens Einschätzungen zur Übertragbarkeit sowie Anpassungsmöglichkeit gegeben werden. Einwände und Verbesserungsvorschläge der einheimischen Akteure konnten durch den iterativen Mikrozyklus systematisch erhoben und analysiert werden, so dass die hier vorgeschlagenen Anpassungen auf Akzeptanz der Akteure stießen.

Als konkrete Ergebnisse dieser Ex-ante-Evaluation wurden Adaptionsbedarfe auf der kulturellen, berufsbildungsspezifischen sowie programmspezifischen Dimension für den Transfer des Peer-Review-Verfahrens von Deutschland in den chinesischen Kontext identifiziert:

Für eine Übertragung des Peer-Review-Verfahrens nach China ist zuerst eine Koordinationsstelle aufzubauen, die als eine unabhängige Einrichtung das Informationsmanagement zwischen den Schulen und den Peer-Teams übernimmt. Sie kann beispielsweise eine erste Qualitätssicherung der Selbstberichte in der Vorbereitungsphase ermöglichen und die Unterlagen für die Peer-Teams vorbereiten. Die Koordinationsstelle kann zugleich die Qualität der Peer-Teams durch Beratungstätigkeit sichern und eine Schlüsselposition einnehmen, um eine partizipatorische und gleichgestellte Zusammenarbeit und kollegiale Stimmung zwischen den Peer-Teams und den Schulen zu schaffen.

Weiterhin kann für das chinesische Peer-Review-Verfahren überlegt werden, das Peer-Team um Vertretende regionaler Betriebe zu erweitern, um die von der Schule gewünschte betriebliche Perspektive auf die Schulentwicklung einzubeziehen.

Die Ergebnisse der Ex-ante-Evaluation zeigen die Wichtigkeit der Vorbereitung der Akteure auf die partizipatorische Eigenschaft des Verfahrens. Auf der Seite der Peers ist eine Peer-Schulung zur Formulierung von nicht bewertendem Feedback und kollegialem Umgang der Schlüssel für ein erfolgreiches Peer-Review-Verfahren. Auf der Seite der Schule ist ebenso vorab der partizipatorische Ansatz des Verfahrens sowie die Mitgestaltungsmöglichkeit der Schule zu betonen.

Zusätzlich ist es notwendig, von dem linear angeordneten Zeitplan des Peer-Review-Verfahrens abzurücken und einen flexibleren Zeitplan mit Raum für kurzfristige Veränderung zu entwickeln.

Zuletzt ist für die Nachhaltigkeit des Verfahrens eine Integrationsmöglichkeit des Peer-Review-Verfahrens im chinesischen Qualitätsmanagementsystem mit den Akteuren der chinesischen Berufsbildungspolitik zu diskutieren.

Diese Erkenntnisse sind im nächsten Schritt vor dem Hintergrund einer landesweiten Übertragung kritisch zu reflektieren, weil die Pilotstudie einige Limitationen besitzt. Zum Beispiel lassen sich aufgrund der unterschiedlich entwickelten Regionen Chinas und den

teilweise heterogenen ethnischen Zugehörigkeiten aus den Ergebnissen keine Aussagen für ganz China ableiten. Des Weiteren agierte die deutsche Forscherin hier als Peerleitung in allen vier Schulen und birgt die Gefahr der Ergebnisverzerrung aufgrund ihrer Doppelrolle. Zusätzlich ist hier die Gefahr des Ethnozentrismus zu nennen, da die Forscherin sich überwiegend Studien und Konstrukte des westlichen Wissenschaftsverständnisses bedient. Folglich können aus der Ex-ante-Evaluation keine generalisierenden Aussagen getroffen werden.

Die Untersuchung beweist dennoch, dass die Ex-ante-Evaluation zu einer Vorbereitung des Policy-Transfers führt, welcher sich auf systematisch erhobene Daten anstatt auf Spekulationen stützt. Folglich können Fehler während des tatsächlichen Policy-Transfers vermieden werden. Zeitliche sowie finanzielle Ressourcen können entsprechend effizient und nachhaltig eingesetzt werden.

Ein erfolgreicher Policy-Transfer ist allerdings von komplex zusammenhängenden Faktoren abhängig. Neben einer guten Vorbereitung und adäquaten Anpassungen erfordert er ebenso weiche Faktoren wie Schlüsselakteure im Policy-Nehmerland, die vom neuen Programm überzeugt und bereit sind, sich dafür einzusetzen. Vor allem zu Beginn des Transfers ist ein Netzwerk zum Kapazitätsaufbau von besonderer Bedeutung. Ein weiterer wichtiger Punkt ist der Zeitfaktor. Für eine nachhaltige Übertragung fremder Verfahren bedarf es Zeit, um diese mit den bereits existenten Strukturen zu harmonisieren. Die Einführung von neuen Elementen kann zu neuen Handlungen und Reaktionen der Betroffenen und Beteiligten führen. Erst mit der Zeit können neue Verhaltensweisen zu Routinen und Gewohnheiten werden (vgl. STOCKMANN 2013; RESNICK u. a. 2015).

Literatur

- BALZER, Lars: Wie werden Evaluationsprojekte erfolgreich? Landau 2005
- BOSCH, Aida; REICHENBACH, Thomas; SCHMIDT, Gert: Organisationskulturen in Deutschland und China: Typische Problemlagen. In: BÖHN, Dieter u. a. (Hrsg.): Deutsche Unternehmen in China. Märkte, Partner, Strategien. Wiesbaden 2003, S. 179–205
- BUHREN, Claus; GIESKE, Mario; ROLFF, Hans-Günter: Stadtweite Peer-Reviews im Rahmen von UQM mit allen Dortmunder Berufskollegs – ein wissenschaftliches Gutachten. Dortmund 2012 – URL: <https://www.bmbf.de/de/imove-training-made-in-germany-323.html> (Zugriff: 29.01.2019)
- BUICHL, Melanie, WILBERS, Karl: Wirksamkeit externer Evaluation – eine Analyse der Wirksamkeit des European Peer Review im Rahmen der österreichischen Qualitätsinitiative Berufsbildung. In: Berufs- und Wirtschaftspädagogik – online, (2011) 21, S. 1–17
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (BMBF) (Hrsg.): Berufsbildung international. iMove: Training – Made in Germany. 2016 – URL: <https://www.bmbf.de/de/imove-training-made-in-germany-323.html> (Zugriff: 23.07.2020)
- CHEN, Guo-Ming; STAROSTA, William J.: Foundations of intercultural communication. Lanham 2005

- CHEN, Huey-Tsyh: Practical program evaluation: Assessing and improving planning, implementation, and effectiveness. Thousand Oaks 2005
- CLART, Philip: Die Religionen Chinas. Göttingen 2009
- DILGER, Bernadette: Herausforderungen der Gestaltungsforschung in der Wirtschaftspädagogik. In: BRAUKMANN, Ulrich; DILGER, Bernadette; KREMER, Hans-Hugo (Hrsg.): Wirtschaftspädagogische Handlungsfelder. Festschrift für Peter F. E. Sloane zum 60. Geburtstag. Detmold 2014, S. 363–378
- DOLOWITZ, David P.; MARSH, David.: Learning from abroad: The role of policy transfer in contemporary policy-making. In: Governance: An International Journal of Policy and Administration 13 (2000) 1, S. 5–24
- FRAUNHOFER MOEZ (Hrsg.): Treibende und hemmende Faktoren im Berufsbildungsexport aus Sicht deutscher Anbieter. Leipzig 2012
- GEORG, Walter: Vergleichende Berufsbildungsforschung. In: RAUNER, Felix (Hrsg.): Handbuch der Berufsbildungsforschung. Bielefeld 2005, S. 186–193
- GÜNTHNER, Susanne: Diskursstrategien in der interkulturellen Kommunikation. Analysen deutsch-chinesischer Gespräche. Tübingen 1993
- GUTKNECHT-GMEINER, Maria: Externe Evaluierung durch Peer Review. Qualitätssicherung und -entwicklung in der beruflichen Erstausbildung. Wiesbaden 2008
- GUTKNECHT-GMEINER, Maria: Peer review impact analysis report. Analysis of 14 transnational European peer reviews carried out in eight European countries 2006–2009. Final report. Wien 2010
- HALL, Edward Twitchell: Beyond culture. New York 1976
- HALL, Edward Twitchell: The hidden dimension. New York 1966
- HALL, Edward Twitchell: The silent language. New York 1959
- HALL, Edward Twitchell; HALL, Mildred Reed: Understanding cultural differences: Germans, French and Americans. Yarmouth, Maine 1990
- HAO, Yan: The reform and modernization of vocational education and training in China, WZB Discussion Paper, No. SP III 2012-304. Berlin 2012
- HEILMANN, Sebastian: Politik und Staat. In: STAIGER, Brunhild (Hrsg.): Länderbericht China. Geschichte Politik Wirtschaft Gesellschaft Kultur. Darmstadt 2002, S. 67–102
- HERRMANN-PILLATH, Carsten: Unternehmensführung im chinesischen Kulturraum: Elemente und Prozesse. Duisburger Arbeitspapiere zur Ostasienwirtschaft. Duisburg 1997
- HOFSTEDE, Geert; HOFSTEDE, Gert Jan; MINKOV, Michael: Cultures and organizations. Software of the mind. Intercultural cooperation and its importance for survival. New York u. a. 2010
- KELL, Adolf: Organisation, Recht und Finanzierung der Berufsbildung. In: ARNOLD, Rolf; LIPSMEIER, Antonius (Hrsg.): Handbuch der Berufsbildung. Wiesbaden 2006, S. 453–484
- LAMNEK, Siegfried: Qualitative Sozialforschung. Methoden und Technik. Weinheim 1993

- LEWIS, Theodore: The problem of cultural fit – what can we learn from borrowing the German dual system? In: *Compare* 37 (2007) 4, S. 463–477
- LI, Junmin: Das Peer-Review-Verfahren als Instrument der Schulentwicklung. Eine Analyse an berufsbildenden Schulen in Nordrhein-Westfalen. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik* 115 (2019) 1, S. 90–111
- LIU, Bangxiang; JIANG, Dayuan; TROWE, Eberhard: Perspektiven der chinesischen Berufsbildung. In: *Berufsbildung in Wissenschaft und Praxis* 35 (2006) 5, S. 38–42
- MAYRING, Philipp: *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim u. a. 2003
- MINISTRY OF EDUCATION OF THE PEOPLE'S REPUBLIC OF CHINA (MoE) (Hrsg.): Beschluss des Jahres 2003 über den Bildungsstandard der höheren Berufsschulen. Bekanntmachung und Bewertung der Pilotinstitute. [关于确定2003年度高职高专院校人才培养工作水平 评估试点院校的通知] 2003 – URL: http://www.moe.gov.cn/srcsite/A08/s7056/200304/t20030401_124743.html (Zugriff: 29.01.2019)
- MOSSBERGER, Karen; TOLBERT, Caroline J.; STANSBURY, Mary: *Virtual inequality: Beyond the digital divide*. Washington D. C. 2003
- MOSSBERGER, Karen; WOLMAN, Harold: Policy transfer as a form of prospective policy evaluation: Challenges and recommendations. In: *Public Administration Review* 63 (2003) 4, S. 428–440
- RESNICK, Lauren B. u. a.: Innovationen implementieren: Von visionären Modellen zur alltäglichen Praxis. In: DUMONT, Hanna; ISTANCE, David; BENAVIDES, Francisco (Hrsg.): *The Nature of learning – Die Natur des Lernens*. Forschungsergebnisse für die Praxis. Weinheim 2015, S. 268–295
- ROSE, Richard: *Learning from comparative public policy. A practical guide*. London, New York 2005
- SCHIPPERS, Stefan: *Systemberatung zwischen Anspruch und Wirklichkeit. Eine wirkungsorientierte Analyse der „Mubarak-Kohl-Initiative“ in Ägypten*. Marburg 2009
- SHANGHAI MUNICIPAL EDUCATION COMMISSION (Hrsg.): Bekanntgabe zum Shanghai Evaluationsplan. Evaluationsumsetzungsbestimmungen und Durchführungshandreichung [上海市教育委员会关于报送上海市高等职业院校人才培养工作评估规划、实施细则和操作规程的函] 2009 – URL: <http://www.shbangde.com/bumen/gaojyj/jxpg/jxpg03.htm> (Zugriff: 29.01.2019)
- SHENKAR, Oded; RONEN, Simcha: Culture, ideology, or economy: A comparative exploration of work goal importance among managers in Chinese societies. In: PRASAD, S. Benjamin (Hrsg.): *Advances in international comparative management*. 5. Aufl. Greenwich, Connecticut 1990, S. 117–134
- SHI, Hongxia: *Kommunikationsprobleme zwischen deutschen Expatriates und Chinesen in der wirtschaftlichen Zusammenarbeit. Empirische Erfahrungen und Analyse der Einflussfaktoren* (Dissertation). Würzburg 2003

- SILVESTRINI, Stefan: Ex-ante-Evaluation. Ein Planungsansatz für die Entwicklungszusammenarbeit. Sozialwissenschaftliche Evaluationsforschung 10. Münster 2011
- SILVESTRINI, Stefan; STOCKMANN, Reinhard: II. Synthese und Metaevaluierung: Die Studie. 1. Einleitung. In: STOCKMANN, Reinhard; SILVESTRINI, Stefan (Hrsg.): Metaevaluierung Berufsbildung. Ziele, Wirkungen und Erfolgsfaktoren der deutschen Berufsbildungszusammenarbeit. Münster u. a. 2013, S. 25–35
- STOCKMANN, Reinhard: Metaevaluierung. In: STOCKMANN, Reinhard; SILVESTRINI, Stefan (Hrsg.): Metaevaluierung Berufsbildung. Ziele, Wirkungen und Erfolgsfaktoren der deutschen Berufsbildungszusammenarbeit. Münster u. a. 2013, S. 108–177
- STOCKMANN, Reinhard; MEYER, Wolfgang: Chinas Berufsbildung im Wandel. 30 Jahre Entwicklungszusammenarbeit mit der Hanns-Seidel-Stiftung. Münster 2017
- TRIANDIS, Harry C.: Individualism & Collectivism. Boulder, Oxford 1995
- VELDE, Christine: Employers' perceptions of graduate competencies and future trends in higher vocational education in China. In: Journal of Vocational Education and Training, 61 (2009) 1, S. 35–51 – URL: <http://dx.doi.org/10.1080/13636820902819974> (Zugriff: 29.01.2019)
- WITZEL, Andreas: Das problemzentrierte Interview. In: FORUM: QUALITATIVE SOZIALFORSCHUNG 1 (2000) 1, S. 1–9
- WOLPIN, Kenneth I.: Ex ante policy evaluation, structural estimation, and model selection. In: The American Economic Review, (2007) 97(2), S. 48–52
- WULZ, Gabriela, JONACH, Michaela; GRAMLINGER, Franz: Peer-Review in QIBB – Erste Ergebnisse zur Umsetzung von externer Evaluation in österreichischen berufsbildenden Schulen. In: bwp@ Berufs- und Wirtschaftspädagogik – online (2011) 21, S. 1–14 – URL: http://www.bwpat.de/ausgabe21/wulz_et_al_bwpat21.pdf (Zugriff: 29.01.2019)
- YANG, Qing: 论新一轮高职人才培养工作评估实践范式的转型 [Eine neue Diskussionsrunde – Die Umsetzung und Paradigmen der Evaluation von Bildungsarbeiten der höheren Berufsbildung]. In: Baijia Lunyuan (2015) 15, S. 24–27
- YAO, Aiguo: 沿革与反思: 我国高等职业教育质量保障制度建设二十年 [20 Jahre Qualitätsmanagement in der höheren Berufsbildung Chinas]. In: 教育与职业 [Bildung und Beruf] 23 (2008) 8, S. 12–13
- YU, Chun-Sheng; TAYLOR, G. Stephen; TUNG, Wei: A cross-cultural comparison of work goals: The United States, Taiwan, and the People's Republic of China. In: ALON, Ilan (Hrsg.): Chinese culture, organizational behavior, and international business management. Westport, Connecticut, London 2003, S. 169–185
- ZHANG, Qiuling: 我国高职教育质量保障的历史研究 [Historische Forschung zum Qualitätsmanagement der höheren Berufsbildung in China] (1980–2012). In: 职业教育研究 [Berufsbildungsforschung] (2013) 28, S. 17–18
- ZHANG, Weiwei; ZHANG, Jianxing: 高职评估新方案试水一年 [Das Ergebnis der Umsetzung des neuen Evaluationskonzeptes der höheren Berufsbildung nach einem Jahr]. In: Education and Vocation (2009) 9, S. 24–28

Rosemarie Sackmann

► **Beziehungen zwischen Ex-ante-Evaluation und formativer Evaluation. Beispiele aus der wissenschaftlichen Begleitung zum Pilotprojekt „Entwicklung und Erprobung eines Konzepts zur Förderung von Anerkennungskosten“**

Am Beispiel eines Pilotprojektes zur Förderung von Anerkennungskosten wird gezeigt, wie Ex-ante- und formative Evaluation im Rahmen wissenschaftlicher Projektbegleitung zusammenwirken: Die Ex-ante-Evaluation leitet einerseits die prozessbegleitende formative Evaluation an, andererseits wirken Befunde aus der formativen Evaluation auf die Konzeptionierungen der Ex-ante-Evaluation zurück. Die bei der Umsetzung des Pilotprojektes in prozessbegleitenden formativen und summativen Evaluationen gewonnenen Einsichten tragen zur Klärung des Konzeptes des laufenden Projektes und zu einer verbesserten Informationsbasis für die Ex-ante-Evaluation anschließender Förderprojekte bei.

1 Evaluationsprozesse

In der Evaluationsforschung ist es üblich, zwischen Ex-ante-, prozessbegleitender und Ex-post-Evaluation zu unterscheiden. Entsprechend ihren Bezeichnungen werden die drei Evaluationstypen oft verschiedenen Stadien von Maßnahmen oder Programmen zugeordnet, etwa: Planungsphase, Implementierungsphase und Wirkungsphase (vgl. STOCKMANN 2004, S. 5f.). Diese Phasenzuordnung wird auch dadurch befördert, dass verschiedene Forschungsdisziplinen typische Schwerpunkte in Evaluationen setzen. So ist die Ex-ante-Evaluation beispielsweise das bevorzugte Evaluationsinstrument in der Stadtplanung, bei Gesetzgebungsverfahren und bei der Folgenabschätzung von komplexen Programmen in Bereichen

wie Technikentwicklung, Klima- oder Datenschutz (vgl. GUYADEEN/SEASONS 2018; SMISMANS 2015). In den entsprechenden Fachdisziplinen werden häufig Ex-ante-Evaluationen, aber selten prozessbegleitende oder Ex-post-Evaluationen durchgeführt. Disziplinen, die demgegenüber eher auf Steuerungsprozesse fokussiert sind, konzentrieren sich auf prozessbegleitende Management- und Umsetzungsfragen, während beispielsweise aus der Perspektive der Politikwissenschaft Wirkungsanalysen und Kosten-Nutzen-Bewertungen im Zentrum stehen, die in der Regel ex post durchgeführt werden. Solche disziplinären Schwerpunktsetzungen tragen dazu bei, dass die Evaluationstypen als voneinander getrennte Verfahren wahrgenommen werden.

Im Unterschied dazu werden in der jüngeren Evaluationsforschung die Evaluationstypen nicht als verschiedene Ansätze, sondern als Aspekte eines dynamischen Evaluationsprozesses angesehen, der alle Phasen einer Maßnahme oder eines Programms umfasst (vgl. KAUTTO/SIMILÄ 2005; OLIVEIRA/PINHO 2011; MERGAERT/MINTO 2015; SAMSET/CHRISTENSEN 2015; GUYADEEN/SEASONS 2018). Diese Auffassung liegt beispielsweise auch aktuellen Konzepten der Evaluation von EU-Programmen zugrunde (vgl. SMISMANS 2015).¹ An diese neueren Auffassungen von dynamischen Evaluationsprozessen knüpft auch die wissenschaftliche Begleitung des Pilotprojektes zum Anerkennungszuschuss an, die in den folgenden Kapiteln vorgestellt wird.

Eine zweite Unterscheidung zwischen Evaluationsansätzen steht quer zu den drei Evaluationstypen: die Unterscheidung zwischen formativer und summativer Evaluation. Formative und prozessbegleitende Evaluation werden oft gleichgesetzt. Die Bezeichnung „formativ“ (gestaltend) verweist auf die Ausgestaltung von Maßnahmen und betrifft meist die Ebene der Umsetzung von Maßnahmen und Programmen. Formative Evaluation kann aber auch konzeptionell relevante Befunde liefern, denn sie kann Lücken in den Konzepten von Programmen und Maßnahmen und in der Ex-ante-Evaluation aufzeigen. Im Fall der summativen (zusammenfassenden) Evaluation legt die Bezeichnung nahe, dass es dabei um abschließende Bewertungen geht. Dies führt dazu, dass summative Evaluation oft mit Ex-post-Evaluation gleichgesetzt wird. Tatsächlich sind summative Evaluationen auch typisch für Ex-post-Evaluationen, sie können aber grundsätzlich in allen drei Evaluationstypen angewandt werden.² Festzuhalten ist: Formative und summative Evaluation sind nach ihren

-
- 1 Im Ansatz des *Design-Based-Research*, der in der Bildungsforschung aktuell prominent ist, ist die Trennung zwischen Planung und Implementierung weitgehend aufgehoben (vgl. REINMANN 2014). In *Design-Based-Research*-Projekten wird eine Intervention im Umsetzungsprozess in iterativen Wechselschritten zwischen Exploration, Konstruktion und Evaluation geplant, entwickelt und erprobt, wobei die Abfolge dieser Teilprozesse flexibel ist (vgl. dazu das Entwicklungsparadigma der Evaluation bei KROMREY 2001).
 - 2 Stockmann betont, dass die formative und summative Evaluation in allen Projektstadien (und damit in allen Evaluationstypen) vorkommen kann (vgl. STOCKMANN 2004, S. 5f.). Kromrey schließt zwar nicht explizit aus, dass es in der Projektdurchführung zu summativer Evaluation kommt, doch ordnet er im Kontrollparadigma der Evaluation die formative Evaluation der Projektdurchführung und die summative Evaluation der nach Projektende durchzuführenden Ex-post-Evaluation zu (vgl. KROMREY 2001).

Aufgaben zu unterscheiden. Formative Evaluation *begleitet die Umsetzung* von Programmen und Maßnahmen, wobei sie sich an einzelnen Bewertungsindikatoren orientiert und ggf. als notwendig angesehene Anpassungsprozesse unterstützt. Bei summativen Evaluationen liegt der Schwerpunkt hingegen auf der *zusammenfassenden Bewertung* eines Evaluationsgegenstandes (Projektplan, Projektdurchführung, intendierte und nicht intendierte Effekte, Erreichung der Projektziele).

Die Bewertungskriterien, die in Evaluationen zur Anwendung kommen, sind grundsätzlich in Bezug auf die jeweiligen Evaluationsgegenstände zu entwickeln. Zu den gebräuchlichen Bewertungsdimensionen zählen beispielsweise Relevanz, Effektivität und Effizienz. Relevanz betrifft die Frage, ob die Projektziele zum Bedarf passen; Effektivität betrifft die Frage, ob die angestrebten Ziele erreicht werden. Das Kriterium der Effizienz wird in der Regel mit monetären Kosten-Nutzen-Relationen in Verbindung gebracht. Generell können die Bewertungskriterien auf der Ebene der Projektumsetzung und auf der Ebene der Ziele von Maßnahmen und Programmen angewandt werden. Allerdings kann die Bewertung der Effektivität einer Maßnahme oder eines Programms gut *ex post*, in Teilen prozessbegleitend aber kaum *ex ante* vorgenommen werden. *Ex ante* kann im Hinblick auf die Zielerreichung von Maßnahmen und Programmen nur geprüft werden, ob in der Planung Bewertungskriterien der Zielerreichung angemessen berücksichtigt wurden.

Diese knappen Erläuterungen zum Evaluationsprozess zeigen die Vielschichtigkeit des Evaluationskonzeptes auf. In der Praxis gibt es enger und weiter gefasste Evaluationsaufgaben; dabei gewinnen umfassendere, auf alle Aspekte zielende Evaluationen in der Evaluationsforschung an Bedeutung. Die wissenschaftliche Begleitung zum Pilotprojekt zur Entwicklung und Erprobung eines Instrumentes zur Förderung von Anerkennungskosten orientiert sich an einem umfassenden Evaluationskonzept. In den folgenden Kapiteln wird anhand von drei Beispielen aus der wissenschaftlichen Begleitung zum Pilotprojekt ein Einblick in die Unterschiedlichkeit des Zusammenwirkens von Ex-ante- und formativer Evaluation gegeben (siehe Kapitel 2.2 und 2.3). Zuvor jedoch soll das Pilotprojekt vorgestellt werden (Kapitel 2.1).

2 Das Förderinstrument und die Aufgaben seiner wissenschaftlichen Begleitung

Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Pilotprojekt zur Entwicklung und Erprobung eines Konzepts zur Förderung von Anerkennungskosten (kurz: Anerkennungszuschuss) ist ein Baustein in der Gesamtstrategie der Bundesregierung zur Fachkräftesicherung durch Ausschöpfung vorhandener Qualifikationen. Beginnend mit gesetzlichen Regelungen zur beruflichen Anerkennung (zunächst im Regelungsbereich des Bundes, anschließend auch auf Länderebene) wurde seit 2012 ein umfangreiches Instrumentarium zur Unterstützung des Transfers von im Ausland erworbenen Qualifikationen

geschaffen. Der Anerkennungszuschuss wurde 2016 etabliert, um die Kostenhürde von Anerkennungsverfahren zu senken.

Essenzieller Bestandteil des Pilotprojektes ist die durchgängige wissenschaftliche Begleitung der Entwicklung und Erprobung des Förderinstrumentes.

2.1 Die Förderung von Anerkennungskosten im Kontext des Berufsqualifikationsfeststellungsgesetzes

Berufsanerkennung unterstützt die Arbeitsmarktintegration von Personen mit im Ausland erworbenen Berufsqualifikationen. Die Forschung zur Arbeitsmarktintegration von Zuwandernden ergab, dass ein erhebliches Potenzial an Berufsqualifikationen unter Zuwandernden vorhanden ist (vgl. SCHANDOCK/BREMSE 2012), dass Arbeitgeber aber die ausländischen Qualifikationen oft nicht einschätzen können, weshalb durch Berufsanerkennung auch in nicht reglementierten Berufen die Chancen auf adäquate Beschäftigung deutlich verbessert werden (vgl. DAMELANG/ABRAHAM 2016; EKERT u. a. 2017). Monitoringberichte zum Anerkennungsgesetz haben aber aufgezeigt, dass Anerkennungsinteressierte keinen Antrag stellen oder sogar einen gestellten Antrag zurückziehen, sofern Finanzierungsmöglichkeiten fehlen (vgl. BMBF 2016, S. 51f.). Dieses Problem betrifft insbesondere nicht reglementierte Berufe, bei denen der Zugang zur Berufsausübung rechtlich nicht von der Berufsanerkennung abhängt. Durch Verzicht auf Berufsanerkennung werden auf der Individualebene Chancen auf dem Arbeitsmarkt verringert und auf gesellschaftlicher Ebene wird das Qualifikationspotenzial nicht ausgeschöpft.

An diese fast³ bundesweite Problematik knüpft der Anerkennungszuschuss an: Personen mit ausländischen Berufsabschlüssen, geringer Eigenleistungsfähigkeit und dem Interesse, ein Anerkennungsverfahren zu durchlaufen, wird die Möglichkeit eines Zuschusses zu den Kosten des Verfahrens geboten. Mit dem Zuschuss soll die Bereitschaft zur Durchführung von Berufsanerkennungsverfahren insbesondere von Personen mit Qualifikationen in nicht reglementierten Berufen erhöht werden. Dieser Personenkreis hat kaum Zugang zu anderen Fördermöglichkeiten und verfügt oft gleichzeitig über geringe Eigenmittel.

Mit dem Zuschuss werden Personen mit ausländischen Berufsabschlüssen gefördert, wenn sie sich seit mindestens drei Monaten in Deutschland aufhalten und wenn ihre finanziellen Eigenmittel unterhalb bestimmter Werte liegen (26.000 Euro Jahresbruttoeinkommen bei Alleinstehenden, 40.000 Euro bei Paaren). Der Zuschuss ist auf 600 Euro begrenzt. Gefördert werden Kosten, die im Zusammenhang mit Berufsanerkennungsverfahren anfallen.

Interessierte wenden sich an Beratungsstellen (sogenannte zuleitende Stellen), die im Antragsformular mitzeichnen. Der Antrag zum Anerkennungszuschuss wird bei der zentralen Förderstelle (Forschungsinstitut Betriebliche Bildung, f-bb, Büro Chemnitz) eingereicht. Die zentrale Förderstelle prüft die Anträge. Sind alle Fördervoraussetzungen erfüllt, erhalten

3 In Baden-Württemberg, Berlin und Hamburg existierten im Jahr 2016 jeweils Landesprogramme, die u. a. Kosten des Anerkennungsverfahrens förderten; das Baden-Württemberger Programm ist 2018 ausgelaufen.

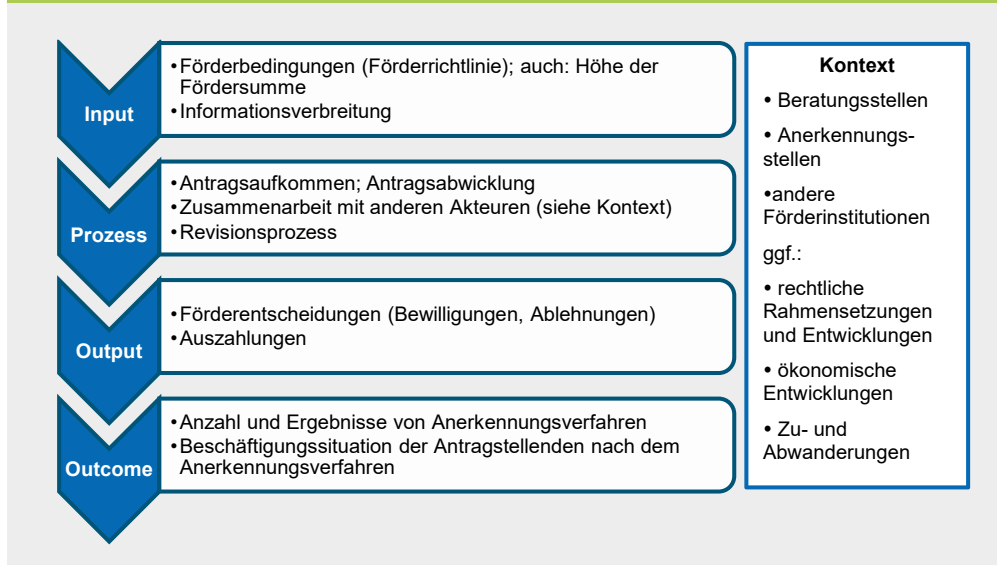
die Antragstellenden eine Förderzusage. Die Erstattung entstandener Kosten erfolgt dann auf Basis eines Weiterleitungsvertrages zwischen der zentralen Förderstelle und den Antragstellenden. Die Auszahlung erfolgt als Erstattung von Ausgaben. Dafür müssen die Geförderten Auszahlungsanträge mit Rechnungen und Zahlungsnachweisen einreichen. Sind für die beantragten Kosten alle Fördervoraussetzungen erfüllt, werden die Kosten an die Antragstellenden erstattet; bei Gebühren, die an die zuständige Stelle im Anerkennungsverfahren zu zahlen sind, ist auch eine Direktüberweisung möglich.

Die zentrale Förderstelle hat den Auftrag, in enger Zusammenarbeit mit dem BMBF das Verfahren zu entwickeln, zu erproben und sämtliche Prozesse, Strukturen, Dokumente und Produkte aufzubereiten und bereitzustellen. Das Pilotprojekt wird vom f-bb durchgehend wissenschaftlich begleitet (Laufzeit des Pilotprojektes: 1. November 2016 bis 30. Juni 2020, verlängert mit neuer Förderrichtlinie seit Oktober 2019 bis Dezember 2022).

2.2 Die wissenschaftliche Begleitung des Projekts

Die wissenschaftliche Begleitung beobachtet sämtliche Aspekte des Pilotprojekts vom Input bis zum Outcome (siehe Abbildung 1).

Abbildung 1: Aufgaben der wissenschaftlichen Begleitung



Der Input des Projektes ist durch die Förderrichtlinie des Anerkennungszuschusses gegeben. Zudem wird hier zum Input auch die Verbreitung von Informationen über das Förderinstrument gerechnet. Die Qualität der Informationsverbreitung wurde von Anfang an beobachtet; auf Schwächen in diesem Bereich wurde umgehend reagiert. Im weiteren Projektverlauf

wurden auch die Förderbedingungen in die projektbegleitenden Bewertungen einbezogen. So wurde beispielsweise geprüft, ob die Fördersumme die Kosten der Anerkennungsverfahren in der Mehrzahl der Förderfälle deckt.

Den Prozessen und dem Output galt in den ersten zwei Jahren des Projektes ein erheblicher Teil der Aufmerksamkeit der wissenschaftlichen Begleitung. Dabei wurden die prozessgenerierten Daten der Förderdatenbank sowie Daten aus Onlinebefragungen der Antragstellenden und der zuleitenden Stellen genutzt. Ab November 2018 trat dann die Analyse des Outcomes in den Vordergrund. Daten zum Outcome werden durch standardisierte Onlinebefragungen der Geförderten generiert.

Besonders hervorzuheben ist hier, dass Kontextfaktoren eine erhebliche Rolle bei der Wirksamkeit des Anerkennungszuschusses spielen. So wirken sich die Beratungsstrukturen auf das Antragsgeschehen aus, da die Beratungsstellen im Anerkennungszuschuss als zuleitende Stellen fungieren. Die zuständigen Stellen von Berufsanzerkennungsverfahren beeinflussen beispielsweise die Höhe der Gebühren von Anerkennungsverfahren. Auch Unterschiede in den rechtlichen Rahmenseetzungen in den Bundesländern wirken auf das Antragsgeschehen. Rechtliche Unterschiede zwischen den Bundesländern verkomplizieren die Antragsprüfung für die zentrale Förderstelle, beispielsweise weil bei Anerkennungsverfahren unterschiedliche Anforderungen im Hinblick auf die einzureichenden Unterlagen gestellt werden. Kosten können nur für Auslagen erstattet werden, die für das Berufsanzerkennungsverfahren notwendig waren. Als besonders aufwendig bei der Antragsbearbeitung hat sich aber die Prüfung anderer Fördermöglichkeiten erwiesen, insbesondere die Möglichkeit der Förderung aus dem Vermittlungsbudget der Agenturen für Arbeit und der Jobcenter: Der Anerkennungszuschuss ist nachrangig zu allen anderen Fördermöglichkeiten. Vor einer Auszahlung ist also zu prüfen, ob ggf. andere Stellen die Kosten erstatten würden (siehe dazu Kapitel 3.3).

Im Rahmen der Ex-ante-Evaluation wurden zu Beginn des Projektes die Ziele und die Indikatoren der Zielerreichung genauer bestimmt (siehe Tabelle 1).

Bei den Zielen ist zwischen der Ebene der Projektumsetzung und der Ebene der Ziele des Förderinstrumentes zu unterscheiden. Auf der Ebene der Projektumsetzung geht es beispielsweise um Effektivität im Hinblick auf die Informationsverbreitung und um Effizienz im Hinblick auf die Verständlichkeit der Formulare und Verfahrensabläufe und die Beratung von Antragstellenden (siehe Tabelle 1). Geprüft wurde die Aufgabenerfüllung teilweise direkt durch Befragungen von Beratungsstellen und Antragstellenden, teilweise anhand indirekter Indikatoren. So kann beispielsweise die Zahl der Anträge und insbesondere der Anteil der positiv entschiedenen Anträge als Hinweis darauf gedeutet werden, dass die Informationen zu den Förderbedingungen gut verbreitet wurden.

Tabelle 1: Zielsetzungen, Indikatoren der Zielerreichung und Datenquellen

Ziele	Indikatoren	Datenquelle
Informationsverbreitung	Anteil Förderzusagen; Bewertung durch zuleitende Stellen	Förderdatenbank Befragung zuleitender Stellen
Verständlichkeit der Formulare und der Verfahrensabläufe	Fehlerquote bei Anträgen Bewertung durch zuleitende Stellen; Bewertung durch die Antragstellenden	Förderdatenbank Befragungen von zuleitenden Stellen, Antragstellenden und Geförderten
Unterstützung der Antragstellenden durch die zentrale Förderstelle	Bewertung durch zuleitende Stellen, Antragstellende und Geförderte	Befragungen von zuleitenden Stellen, Antragstellenden und Geförderten
Zügige Verfahrensabwicklung	Dauer der Antragsbearbeitung	Förderdatenbank
Angemessenheit der Höchstfördersumme	Höhe der ausgezahlten Förderungen Bewertung durch zuleitende Stellen Angaben zu Anerkennungskosten seitens der Geförderten	Förderdatenbank Befragungen zuleitender Stellen Gefördertenbefragungen
Projekt-Output	Zahl der Anträge Erreichen der Zielgruppe (insbesondere: nicht reglementierte Berufe, erwerbstätige Antragstellende)	Förderdatenbank
Bewertung des Nutzens des Förderinstrumentes	Einschätzungen zuleitender Stellen und Geförderter	Quantitative und qualitative Befragungen zuleitender Stellen Onlinebefragung Geförderter
Projekt-Outcome	Durchführung von Berufsamerkenungsverfahren Ergebnisse der Berufsamerkenungsverfahren: <ul style="list-style-type: none"> ▶ Teilnahme an Qualifizierungsmaßnahmen ▶ aus der Arbeitslosigkeit in Beschäftigung ▶ aus dem Minijob in reguläre Beschäftigung ▶ aus einer qualifikationsinadäquaten Beschäftigung in eine adäquate[re] ▶ aus einer befristeten in eine unbefristete Beschäftigung ▶ Einkommenssteigerung 	Befragung der Geförderten

Auf der Ebene der Ziele des Förderinstrumentes geht es einerseits um das Kriterium der Relevanz, also um die Frage, ob der Anerkennungszuschuss den Bedarf trifft. Als Indikatoren können hier die Antragszahlen dienen sowie auch Bewertungen durch Antragstellende und zuleitende Stellen. Andererseits geht es um die Bewertungsdimension der Effektivität und damit um die Frage, ob die Zielgruppe des Förderinstrumentes erreicht wird. Bei den entsprechenden Untersuchungen zum Output des Projektes handelt es sich um typische Aufga-

ben des Monitorings (siehe dazu auch Kapitel 3.1).⁴ Die Untersuchung des Projekt-Outcome ist dagegen als Ex-post-Evaluation einzuordnen: Untersucht wird, wie es für die Geförderten nach Abschluss der Förderung weiter ging. Zentrale Fragen dabei sind, ob und mit welchen Ergebnissen die Geförderten ein Anerkennungsverfahren durchlaufen und ob sich in der Folge ihre Erwerbssituation positiv verändert.

Die wissenschaftliche Begleitung zum Pilotprojekt führt also Ex-ante-, prozessbegleitende und Ex-post-Evaluationen durch. Und sie nutzt dabei formative wie auch summative Evaluationsansätze. Im folgenden Abschnitt wird näher betrachtet, in welcher Relation diese verschiedenen Aufgaben der wissenschaftlichen Begleitung des Pilotprojektes zueinander stehen.

2.3 Monitoring und Evaluation in der wissenschaftlichen Begleitung

Viele der oben beschriebenen Aufgaben der wissenschaftlichen Begleitung im Anerkennungszuspruch sind dem Monitoring zuzuordnen, weil es sich dabei um wiederholte Beobachtungen handelt, die der Kontrolle der Erreichung der Projektziele dienen (vgl. GEBEL 2001; ähnlich STOCKMANN 2004, S. 9f.).⁵ Die Aufgaben der wissenschaftlichen Begleitung im Pilotprojekt gehen aber über typische Monitoringtätigkeiten hinaus: Die Projektdaten werden systematisch auf konzeptionell relevante Befunde hin analysiert, und es werden Zusammenhangs- und Wirkungsanalysen durchgeführt, durch deren Ergebnisse Annahmen aus der Ex-ante-Evaluation präzisiert und ergänzt werden. In Abbildung 2 werden die Beziehungen zwischen den Monitoringaufgaben (schwarze Schrift, schwarze Linien) und weitergehenden Aufgaben der wissenschaftlichen Begleitung (blaue Schrift und blaue Linien) dargestellt. Die Abbildung soll die Kreisläufe der Wissensgenerierung zwischen Ex-ante-Evaluation und Wirkungsanalyse veranschaulichen. Nicht in jedem Fall läuft der Prozess der Evaluation über die formative Evaluation, die in der Mitte der Abbildung steht. Doch insbesondere in der ersten Zeit der Projektumsetzung hat die wissenschaftliche Projektbegleitung gerade aus der Verbindung zur formativen Evaluation wichtige Impulse gewonnen.

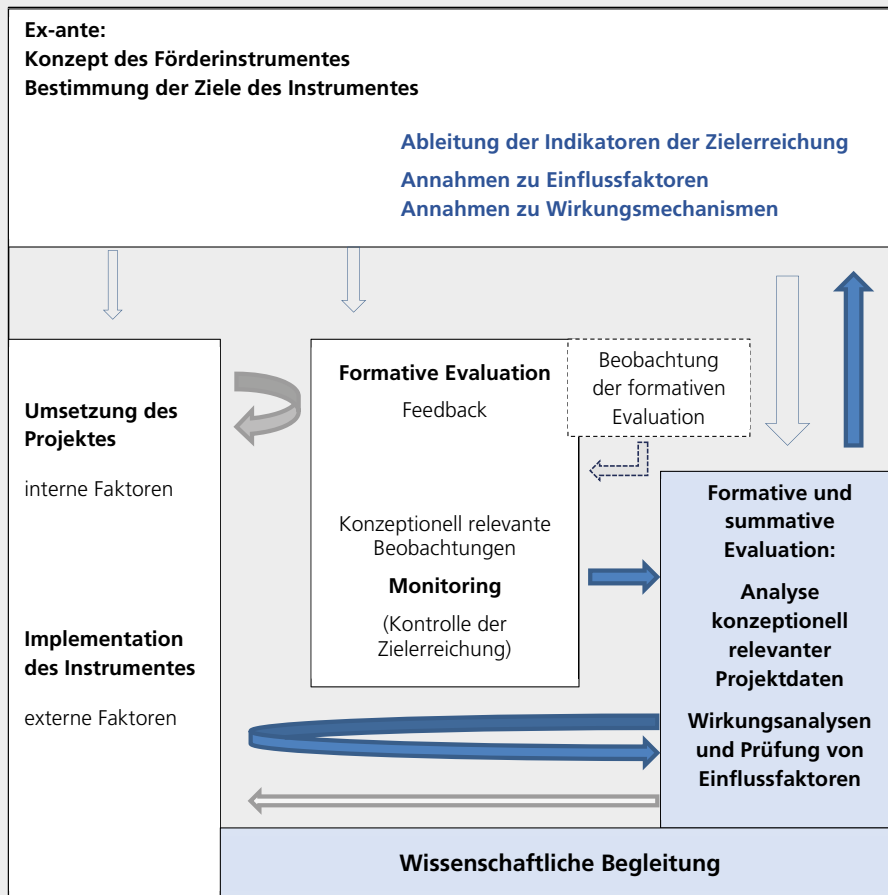
In den ersten Monaten des Pilotprojektes erfolgte primär eine direkte Begleitung nah an der Projektumsetzung. Auf der Basis von Reaktionen von Antragstellenden und zuleitenden Stellen wurde in kurzen Abständen regelmäßig sowie bei gegebenem Anlass Feedback in die Projektentwicklung eingespeist. Diese formative Evaluation blieb im Pilotprojekt Bestandteil des operativen Geschehens. Sie wurde zwar dokumentiert, die aufgetretenen Probleme und

4 Dabei ist auch zu prüfen, ob die Zusammensetzung der Antragstellenden im Hinblick auf das Geschlecht ausgewogen und in Hinblick auf den Wohnort proportional angemessen ist.

5 Ein alternativer Vorschlag besteht darin, die wissenschaftliche Begleitung zum Anerkennungszuspruch insgesamt als Ex-ante-Evaluation zu begreifen, da es sich um ein Pilotprojekt handelt, dessen Aufgabe die Entwicklung und Erprobung eines Förderinstrumentes ist. Auf diese Möglichkeit der Einordnung hat Reinhard Stockmann im Rahmen des AG BFN-Workshops „Evaluation und Wirkungsforschung in der Berufsbildung“ (November 2018) hingewiesen (vgl. McDAVID/HAWTHORN 2006, S. 22).

die daran anschließenden Aktionen und Reaktionen wurden aber keiner weiteren Analyse unterzogen.⁶

Abbildung 2: Das Zusammenwirken der verschiedenen Aufgaben der wissenschaftlichen Begleitung



Quelle: eigene Darstellung

⁶ Man kann hier die Frage aufwerfen, ob es sinnvoll ist, diese Art von Feedbackspeisung als Evaluation zu bezeichnen. Bislang werden jedoch direkt in Projekte eingebundene formative Evaluationen sehr selten einer höherstufigen Analyse unterzogen. Diese Reflexionsebene wird in vielen Projekten für unnötig gehalten, manchmal wird sie wohl auch als eine bedrohliche Form der Kontrolle empfunden. Das Potenzial zur Generalisierung von Problemlösungsstrategien wird so meist nicht genutzt. Andererseits sind wissenschaftliche Begleitungen so vielfältig ausgerichtet, dass die Verfolgung weiterer Aufgaben auch zu einer Überforderung führen könnte.

Das Monitoring zum Projekt beobachtet vor allem den Output. Regelmäßig ausgewertet und in Berichten dargestellt werden beispielsweise Zahlen zu den Antragstellenden (u. a. im Hinblick auf den Referenzberuf und das Land des Berufsabschlusses) und Angaben zu den Fördersummen und Fördergegenständen. Die primäre Datenbasis für das Monitoring ist die Förderdatenbank des Anerkennungszuschusses. Neben den prozessbegleitenden Monitoringaufgaben werden in der wissenschaftlichen Begleitung zum Anerkennungszuschuss Befragungen von zuleitenden Stellen und von Antragstellenden durchgeführt. Von besonderer Bedeutung für die Analysen des Outcome sind Verbleibsbefragungen der Geförderten. Das primäre Ziel der Verbleibsbefragungen ist die Datengenerierung für Wirkungsanalysen.

In Verbindung mit den Datendokumentationen und den systematischen Datenanalysen kommt es im Projektverlauf zu geplanten wie auch zu ungeplanten konzeptionell relevanten Beobachtungen. In Kapitel 3 werden die Beziehungen zwischen Ex-ante- und formativer Evaluation in der wissenschaftlichen Begleitung zum Anerkennungszuschuss mit Beispielen veranschaulicht.

3 Beziehungen zwischen Ex-ante- und formativer Evaluation

Die Beziehungen zwischen Ex-ante-Evaluation und formativer Evaluation sind in der wissenschaftlichen Begleitung zum Anerkennungszuschuss vielfältig. Dies wird in den folgenden Abschnitten anhand von Beispielen näher dargelegt.

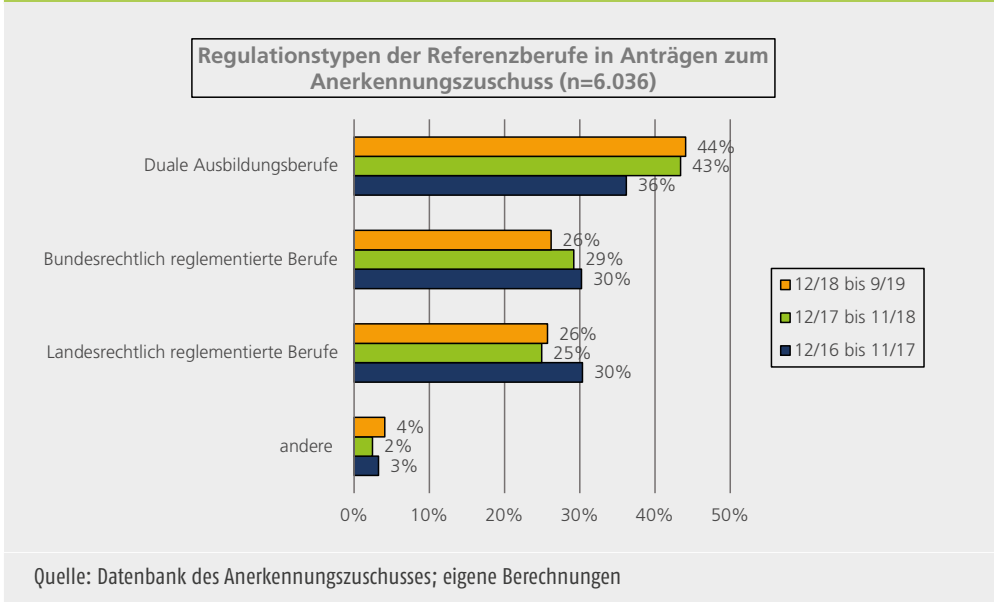
3.1 Formative Evaluation auf Basis von ex ante definierten Zielen

In der Ex-ante-Evaluation zum Pilotprojekt war (u. a.) die Zielgruppe für den Zuschuss näher bestimmt worden.⁷ In der Förderrichtlinie wird (u. a.) besonderes Gewicht auf die in Anerkennungsverfahren bislang unterrepräsentierten nicht reglementierten Berufe gelegt. Abbildung 3 zeigt die Anteile der verschiedenen Regulationstypen an den Anträgen zum Anerkennungszuschuss zwischen Dezember 2016 und September 2019.⁸

7 Überwiegend erfolgte die Bestimmung der Zielgruppe anhand der Förderrichtlinie. Darüber hinaus sollte die Förderung durch ein Bundesprogramm erfolgen, aber weder regionale Verzerrungen noch geschlechtsbezogene Ungleichverteilungen aufweisen. Beides ist faktisch auch nicht der Fall.

8 Die Daten werden hier bis 30. September 2019 dargestellt, weil bis zu diesem Datum die Antragstellung nach der ersten Förderrichtlinie des Anerkennungszuschusses möglich war. Ab Oktober 2019 wurde der Zuschuss auf der Basis einer neuen Förderrichtlinie weitergeführt.

Abbildung 3: Verteilung der Anträge zum Anerkennungszuspruch nach Regulationstyp der Referenzberufe



Ein Zielwert für den Anteil von nicht reglementierten Referenzberufen war ex ante nicht bestimmt worden. Für die Evaluation des Output werden daher Angaben aus dem Monitoring der Berufsanerkennung für bundesrechtlich geregelte Berufe als Vergleichsdaten genutzt: Von 2012 bis 2017 lag der Anteil der nicht reglementierten Berufe in dieser Statistik zwischen 23 und 25 Prozent (vgl. BMBF 2019, S. 29; SCHMITZ 2018). Dagegen lag der Anteil der nicht reglementierten Referenzberufe, d. h. der dualen Ausbildungsberufe und der nicht reglementierten Fortbildungsberufe, unter den Anträgen zum Anerkennungszuspruch zwischen Dezember 2016 und November 2017 bereits bei 36 Prozent. Im zweiten Jahr der Förderung ist der Anteil auf 43 Prozent gestiegen, und im dritten Jahr lag er bei 44 Prozent. Der vergleichsweise hohe Anteil nicht reglementierter Berufe im Anerkennungszuspruch kann als Beleg für die Erreichung der Zielgruppe gewertet werden.

Die Verteilung der Anträge zum Anerkennungszuspruch nach Regulationstyp wird hier als ein Beispiel für Monitoringaufgaben der wissenschaftlichen Begleitung angeführt. Die Beziehung zwischen Ex-ante- und formativer Evaluation ist dabei einfach: Durch die Ex-ante-Evaluation in der Projektkonzeption wurde vorgegeben, worauf in der fortlaufenden formativen Evaluation des Projektes zu achten ist. Wenn die Evaluation wie im Beispiel zu der Einschätzung kommt, dass die Zielsetzung erreicht wird, sind in dieser Hinsicht weder Eingriffe noch weitergehende Untersuchungen nötig.

3.2 Formative Evaluation zur Korrektur der Ex-ante-Evaluation

Generell kann angenommen werden, dass im Rahmen der umfassenden Evaluationsprozesse von wissenschaftlichen Begleitungen Erkenntnisse gewonnen werden, von denen Rückwirkungen auf vorab formulierte Annahmen und Zieldefinitionen ausgehen können. Auf diese Weise kann die formative Evaluation ergänzend auf die Ex-ante-Evaluation zurückwirken. Im Folgenden soll ein einfaches Beispiel für die rückwirkende Ergänzung der für die Evaluation vorab getroffenen Festlegungen etwas ausführlicher dargestellt werden: die Festlegung des Personenkreises für die Verbleibsbefragungen.

Ursprünglich wurde davon ausgegangen, dass die Verbleibsbefragungen sechs Monate nach Förderabschluss erfolgen sollten. Der Abschluss eines Förderfalles tritt durch Ausschöpfen der Fördersumme ein oder dadurch, dass die Geförderten mitteilen, dass sie keine weiteren Auszahlungen beantragen werden. Davon ausgehend hätten für die erste Welle der Verbleibsbefragungen Anfang November 2018 maximal 200 Personen angeschrieben werden können.⁹ Da von einer Rücklaufquote von rund 33 Prozent ausgegangen wurde, wäre die Datenbasis für Analysen zu diesem Zeitpunkt sehr schmal gewesen. Durch das Monitoring der Auszahlungen war allerdings bekannt, dass für über 800 Geförderte bereits Zahlungen für Gebühren von Anerkennungsverfahren erfolgt waren. Zudem legte die Beobachtung von Aktivitäten bzw. von Nicht-Aktivitäten der Geförderten die Vermutung nahe, dass viele darauf verzichteten, gegenüber der zentralen Förderstelle das Ende ihres Förderbedarfs anzuzeigen. Daher wurde die ursprüngliche Festlegung revidiert: In die Verbleibsbefragungen werden nun auch Personen eingeschlossen, bei denen mindestens sechs Monate vor dem Befragungszeitpunkt Gebühren an eine zuständige Stelle für Berufsanerkennungsverfahren gezahlt wurden. Damit konnten in den ersten beiden Befragungswellen (November 2018 und Mai 2019) 857 Geförderte angeschrieben werden. Die Nettorücklaufquote betrug 50 Prozent, sodass mit den Daten der ersten Befragungswellen bereits informative Analysen durchgeführt werden konnten.¹⁰

Das im Monitoring gesammelte Wissen über das Antragsgeschehen hat in diesem Beispiel eine sinnvolle Abänderung der ex ante festgelegten Verfahrensweise des Monitorings ermöglicht. Zudem wurde durch die parallel zur Projektumsetzung durchgeführten Outcome-Analysen beispielsweise deutlich, dass die Unterscheidung zwischen reglementierten und nicht reglementierten Berufen für Wirkungsanalysen differenziert werden sollte, da die Unterschiede zwischen den Berufen erheblich sein können.¹¹

9 Von dieser Zahl gehen noch diejenigen Geförderten ab, die eine Befragung im Rahmen der wissenschaftlichen Begleitung bereits im Antrag auf den Anerkennungszuschuss generell abgelehnt haben.

10 Inzwischen (August 2020) sind vier Wellen der Verbleibsbefragungen durchgeführt worden. Der Nettorücklauf liegt weiter bei 50 Prozent (Stand August 2020: 1.602 Befragte und 806 auswertbare Antworten).

11 Die Ergebnisse der projektbegleitend durchgeführten Verbleibsbefragungen haben auch einen Beitrag zur Konzipierung eines neuen Instruments zur Förderung von Qualifizierungskosten geleistet. Das Pilotprojekt dazu ist im Januar 2020 gestartet.

3.3 Formative Evaluation als Vorlage für Ex-ante-Evaluationen in Folgeprojekten

Allgemein gilt es als wünschenswert, dass Ex-post-Evaluationen strukturgebend für Folgeprojekte werden (vgl. MERGAERT/MINTO 2015). Formative Evaluation wird in solchen Zusammenhängen in der Regel nicht erwähnt. Die Vorstellung der Bewertung von Projektergebnissen verbindet sich quasi automatisch mit (oft extern durchgeführten) Ex-post-Evaluationen. Im Folgenden wird an einem Beispiel gezeigt, dass bereits aus der prozessbegleitenden Evaluation Ergebnisse hervorgehen können, die auch für die Ex-ante-Evaluation von Folgeprojekten relevant sind.

Der Anerkennungszuschuss ist nachrangig zu allen anderen Fördermöglichkeiten. Bevor Fördermittel aus dem Zuschuss gezahlt werden können, muss daher geklärt werden, ob die Förderung durch andere Institutionen erfolgen könnte. Anerkennungskosten könnten in drei Bundesländern (Berlin, Baden-Württemberg, Hamburg) über Landesprogramme gefördert werden. Mit diesen Institutionen konnten beim Start des Anerkennungszuschusses tragfähige Absprachen über die gegenseitige Information zu Antragsfällen getroffen werden. Bei arbeitslosen Anerkennungsinteressierten ist zudem eine Förderung von Anerkennungskosten durch Jobcenter oder Arbeitsagenturen möglich. Ob diese erfolgt, hängt vom Ermessen der Fallbetreuung ab; begrenzend wirken dabei generell die verfügbaren Ressourcen der jeweiligen Einrichtung. Wie bei den Landesprogrammen *könnte* also eine Förderung durch Jobcenter oder Arbeitsagentur erfolgen. Anders als bei den Landesprogrammen handelt es sich aber bei den Jobcentern und Arbeitsagenturen um eine Vielzahl an Unterorganisationen der Bundesagentur für Arbeit oder auch um kommunale Einrichtungen. An der Schnittstelle des Anerkennungszuschusses zu Jobcentern und Arbeitsagenturen sind bei der Umsetzung des Zuschusses Abstimmungsprobleme aufgetreten.

Die Grundkonstellation ist folgende: Um sicherzustellen, dass Antragstellende beim Anerkennungszuschuss nicht auch aus den Mitteln der Arbeitsagenturen und Jobcenter gefördert werden können (oder sogar bereits Förderung für ihr Anerkennungsverfahren erhalten), ist eine Anlage zum Antrag auf Anerkennungszuschuss entwickelt worden („Anlage B“). Die Anlage ist von der zuständigen Agentur oder dem zuständigen Jobcenter auszufüllen. Diese Anlage müssen alle Antragstellenden einreichen, die arbeitslos, arbeitssuchend oder geringfügig beschäftigt sind.

Probleme mit der Anlage B wurden erstmals sichtbar, als die zentrale Förderstelle durch Mitteilungen von zuleitenden Stellen und Antragstellenden darauf aufmerksam wurde, dass einige Sachbearbeiter/-innen in Arbeitsagenturen und Jobcentern sich weigerten, die Anlage B auszufüllen. Durch Absprachen mit der Bundesagentur für Arbeit (BA) wurde versucht, dieses Problem zu lösen. Es sollte eine Handlungsempfehlung und Geschäftsanweisung der BA (HEGA) zu dem Thema herausgegeben werden. Die Veröffentlichung verzögerte sich jedoch und in der Zwischenzeit entwickelte sich das Problem weiter: In Agenturen und Jobcentern wurde eine E-Mail verbreitet, die so formuliert war, dass viele Empfänger/-innen daraus ableiteten, dass der Anerkennungszuschuss Vorrang vor der Förderung durch die Agenturen und Jobcenter habe. Informationen darüber wurden der zentralen Förderstelle direkt

von Personen aus Agenturen und Jobcentern und durch zuleitende Stellen mitgeteilt. Damit verschärften sich die Probleme für Anträge mit Anlage B: Anfangs hatte es genügt, wenn Arbeitsagenturen oder Jobcenter in der Anlage B angaben, dass sie nicht fördern würden. Da aber nicht ausgeschlossen werden konnte, dass Arbeitsagenturen und Jobcenter nicht förderten, weil sie den Zuschuss fälschlich als vorrangig eingestuft hatten, genügte der zentralen Förderstelle die Mitteilung über Nicht-Förderung allein nicht mehr. Es muss seitdem auch ein Grund für die Nichtförderung angegeben sein, damit ausgeschlossen werden kann, dass die Nichtförderung auf falschen Vorstellungen von der Vor- bzw. Nachrangigkeit der Förderungen beruht.

Mit der Veröffentlichung der HEGA der BA (Weisung 201705011 vom 22. Mai 2017) wurde das Ausgangsproblem deutlich gemildert, doch das Folgeproblem wurde damit nicht gelöst: Während anfangs das Problem in der Beschaffung der Anlage B lag, war nun die Begründung der Nicht-Förderung zum größten Problem für die Antragsbearbeitung geworden.

Über einen längeren Zeitraum wurde in der zentralen Förderstelle versucht, in jedem Einzelfall die Beachtung der Nachrangigkeit des Zuschusses zu anderen Fördermöglichkeiten sicher zu stellen und so der Sorgfaltspflicht zu genügen. Dies hatte allerdings erhebliche Folgen für eine andere Anforderung an die Antragsbearbeitung: Bearbeitungszeiten sollten möglichst kurz sein. Die Bearbeitungszeiten wurden in Fällen mit Anlage B aber oft deutlich länger als bei anderen Anträgen. Bei Anträgen, für die Anlage B erforderlich ist, kommt es häufig zu Verzögerungen, weil die Anlage nicht mit dem Antrag eingereicht wird und nachgefordert werden muss. Wie lange es bis zur Nachlieferung dauert, hängt davon ab, wie schnell die Antragstellenden einen Termin bei der Arbeitsagentur oder dem Jobcenter bekommen und wie problemlos die Beschaffung der Anlage verläuft. Bei Anträgen, bei denen seitens der zentralen Förderstelle eine Nachfrage zu den Gründen für die Nicht-Förderung erfolgen muss, kann sich die Bearbeitungsdauer der Anträge nochmals erheblich verlängern, weil die Ansprechpartner/-innen nicht erreichbar sind oder auf die Nachfragen nicht reagieren.¹² In Tabelle 2 sind die unterschiedlichen Bearbeitungsdauern dargestellt. Erfasst sind dabei Anträge, bei denen die Prüfung zum Zeitpunkt der Datenerhebung abgeschlossen war.

12 Bei den Anträgen, bei denen die Anlage B und die Klärung von Gründen für Nicht-Förderung erforderlich sind, lagen im November 2018 durchschnittlich 60 Tage Bearbeitungszeit zwischen der Nachfrage bei der Arbeitsagentur oder dem Jobcenter bis zur Beantwortung der Nachfrage. Inzwischen wurde in Absprache mit dem BMBF für diesen Teil der Probleme eine Regelung gefunden, durch die der Bearbeitungsaufwand und die Bearbeitungsdauer deutlich verringert werden. Das grundlegende Problem mit der Anlage B bleibt davon jedoch unberührt.

Tabelle 2: Bearbeitungsdauer zwischen Antragseingang und Förderzusage

	Bearbeitungsdauer Mittelwert (in Tagen)	Bearbeitungsdauer Median (in Tagen)
Anträge ohne Anlage B	51	40
Anträge mit Anlage B	82	64
Anträge mit Anlage B und Nachfrage bei Arbeitsagenturen oder Jobcentern zu Gründen für die Nicht-Förderung	125	108

Quelle: Datenbank der Anträge zum Anerkennungszuspruch; eigene Berechnungen (Datenbasis: 2.264 Fälle mit Förderzusage; Stand: 08.11.2018)

Die Daten zur Bearbeitungsdauer zeigen, dass die Art und Weise, wie die Beachtung der Nachrangigkeit gesichert wurde, zu differenziellen Effekten unter den Antragstellenden geführt hat. Für Antragstellende, die die Anlage B einreichen mussten, war nicht nur der Aufwand für die Antragstellung größer, sie mussten durchschnittlich auch deutlich länger auf die Förderzusage warten. Für sie verlängerte sich die Bearbeitungsdauer durchschnittlich um 61 Prozent (Median: 60 %). Wo eine Nachfrage zu den Gründen für die Nicht-Förderung durch Agenturen für Arbeit oder Jobcenter nötig war, verlängerte sich die Bearbeitungsdauer durchschnittlich um 145 Prozent (Median: 110 %). Diese Bearbeitungszeiten schieben auch die Eröffnung eines Berufsanerkenntnisverfahrens deutlich hinaus.

Das geschilderte Problem mit der Sicherung der Nachrangigkeit des Anerkennungszuspruches war in der Ex-ante-Evaluation des Pilotprojektes nicht vorhergesehen worden; und da die größten Schwierigkeiten aus der Verbreitung von fehlerhaften Informationen hervorgingen, konnten die Schwierigkeiten so auch nicht vorhergesehen werden.¹³ Der Versuch, in jedem Fall die Beachtung der Nachrangigkeit zu belegen, hat allerdings den Arbeitsaufwand für die zentrale Förderstelle stark erhöht und zu erheblichen Frustrationen bei den Mitarbeitern und Mitarbeiterinnen (sowie vermutlich auch bei einigen Sachbearbeitern in Jobcentern und Arbeitsagenturen) geführt. Gelöst wurde das Problem erst, als im Rahmen der formativen Evaluation die Aufmerksamkeit auf die Verlängerung der Wartezeiten und die Auswirkungen auf die Dauer des Anerkennungsprozesses der Antragstellenden gelenkt wurde. Damit wurde neben der Sorgfaltspflicht das Kriterium der Verfahrensdauer stärker sichtbar. Im Anschluss wurde eine allgemeine Regel zur Bearbeitung der Anträge gefunden, bei denen Nachfragen bei Arbeitsagenturen und Jobcentern wiederholt ohne Erfolg blieben. Mit dieser Regel konnten die beschriebenen Probleme gelöst werden.

Aus dem Beispiel kann man für Ex-ante-Evaluationen ähnlicher Förderinstrumente lernen, dass es sinnvoll ist, bei einer Vielzahl an Kooperationspartnern (wie beispielsweise hunderten von Jobcentern und Arbeitsagenturen und deren einzelnen Sachbearbeitern und

13 Es ist nicht ungewöhnlich, sondern eher die Regel, dass Ex-ante-Evaluationen Lücken aufweisen (vgl. SAMSET/CHRISTENSEN 2017).

Sachbearbeiterinnen) eine allgemeine und einfach handhabbare Regel für die Verfahrensweise bei Kooperationsfehlschlägen zu formulieren.

4 Vielfalt der Beziehungen zwischen Ex-ante- und formativer Evaluation

Ziel des vorliegenden Beitrags war es aufzuzeigen, dass es zwischen Ex-ante- und formativer Evaluation vielfältige Beziehungen gibt. An Beispielen aus der wissenschaftlichen Begleitung zum Anerkennungszuschuss wurden drei Beziehungen dargestellt:

► **Die formative Evaluation orientiert sich an den Vorgaben der Ex-ante-Evaluation (siehe oben Kapitel 3.1).**

Die Festlegungen von Evaluationsgegenständen, die in der Ex-ante-Evaluation vorgenommen wurden, strukturieren die prozessbegleitende Evaluation. Auch wenn sich aus den prozessbegleitenden Beobachtungen keine Handlungserfordernisse ergeben, handelt es sich um formative Evaluation, weil sie nicht auf zusammenfassende, abschließende Bewertungen, sondern auf die Begleitung der Umsetzung ausgerichtet ist.

► **Die formative Evaluation erbringt Befunde, die zur Korrektur von Vorgaben der Ex-ante-Evaluation im laufenden Projekt führen. Die formative Evaluation wirkt also auf die Ex-ante-Evaluation zurück (siehe oben Kapitel 3.2).**

Die Korrektur von ex ante formulierten Festlegungen durch Befunde der formativen Evaluation, umfasst Anpassungen in der Projektumsetzung. Es geht dabei **nicht** um Veränderungen der Projektkonzeption.

► **Die formative Evaluation legt Sachverhalte offen, die für die Ex-ante-Evaluation künftiger Projekte relevant sind (siehe oben Kapitel 3.3).**

Die Beobachtung von Durchführungsproblemen im Rahmen der formativen Evaluation hat im Beispiel zu Analysen geführt, die über die formative Ebene der Projektumsetzung hinausgehen. Aufgrund von Kontexteinflüssen produzieren die Umsetzungsprozesse unbeabsichtigte differenzielle Effekte.

Die dargestellten Beispiele zeigen, dass die wissenschaftliche Begleitung von Pilotprojekten dem in der neueren Forschung verbreiteten Konzept von Evaluation entspricht: Evaluation wird als Prozess verstanden, bei dem die formative Evaluation und die Ex-ante-Evaluation (wie auch mit der Ex-post-Evaluation) in Austauschbeziehungen miteinander stehen und sich so gegenseitig befördern. Eine intensive wissenschaftliche Begleitung kann den Erkenntnisgewinn von Pilotprojekten deutlich steigern. Auch die Chancen, dass aus Evaluationen frühzeitig Informationen für Folgeprojekte hervorgehen, werden vergrößert. Der umfassende Ansatz bietet somit einige Vorteile gegenüber reinen Ex-post-Evaluationen.

Allerdings birgt die enge Verbindung zwischen Projekt und wissenschaftlicher Begleitung die Gefahr, dass die Kritikfähigkeit der Evaluation eingeschränkt wird. Diesem Problem

kann durch Selbstreflexion und Selbstkontrolle im Verlauf des Projektes begegnet werden, doch letztlich bleibt die externe Evaluation ein zentrales Mittel der Erfolgskontrolle.

Literatur

- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (BMBF) (Hrsg.): Bericht zum Anerkennungsgesetz 2016. Berlin 2016 – URL: https://www.bmbf.de/upload_filestore/pub/Bericht_zum_Anerkennungsgesetz_2016.pdf (Zugriff: 24.07.2020)
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (BMBF) (Hrsg.): Bericht zum Anerkennungsgesetz 2019. Berlin 2017 – URL: https://www.bmbf.de/upload_filestore/pub/Bericht_zum_Anerkennungsgesetz_2019.pdf (Zugriff: 22.08.2020)
- DAMELANG, Andreas; ABRAHAM, Martin: You Can Take Some of It with You! A Vignette Study on the Acceptance of Foreign Vocational Certificates and Ethnic Inequality in the German Labor Market. In: Zeitschrift für Soziologie 45 (2016) 2, S. 91–106 – URL: <https://www.degruyter.com/document/doi/10.1515/zfsoz-2015-1005/html> (Zugriff: 09.03.2021)
- EKERT, Stefan u. a.: Evaluation des Anerkennungsgesetzes. Endbericht. Berlin, Frankfurt am Main 2017 – URL: https://www.anerkennung-in-deutschland.de/assets/content/Medien_Dokumente-AI/evaluationsbericht.pdf (Zugriff: 09.03.2021)
- GEBEL, Michael: Monitoring und Benchmarking bei arbeitsmarktpolitischen Maßnahmen. Dokumentation Nr. 06-01, Zentrum für Europäische Wirtschaftsforschung. Mannheim 2001
- GUYADEEN, Dave; SEASONS, Mark: Evaluation Theory and Practice: Comparing Program Evaluation and Evaluation in Planning. In: Journal of Planning Education and Research 38 (2018) 1, S. 98–110
- KAUTTO, Petrus; SIMILÄ, Jukka: Recently Introduced Policy Instruments and Intervention Theories. In: Evaluation 11 (2005) 1, S. 55–68
- KROMREY, Helmut: Evaluation – ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis. In: Sozialwissenschaften und Berufspraxis (2001) 2, S. 1–23
- MCDONALD, James C.; HAWTHORNE, Laura R. L.: Program Evaluation & Performance Measurement: An Introduction to Practice. Thousand Oaks, London, New Delhi 2006
- MERGAERT, Lut; MINTO, Rachel: Ex ante and Ex Post Evaluations: Two Sides of the Same Coin? The Case of Gender Mainstreaming in EU Research Policy. In: European Journal of Risk Regulation (2015) 1, S. 47–56
- OLIVEIRA, Vitor; PINHO, Paulo: Bridging the Gap Between Planning Evaluation and Programme Evaluation: The Contribution of the PPR methodology. In: Evaluation 17 (2011) 3, S. 293–307
- REINMANN, Gabriele: Welchen Stellenwert hat die Entwicklung im Kontext von Design Research? Wie wird Entwicklung zu einem wissenschaftlichen Akt? In: EULER, Dieter;

- SLOANE, Peter F. E. (Hrsg.): Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft „Design-based Research“ (2014), S. 63–78
- SAMSET, Knut; CHRISTENSEN, Tom: Ex Ante Project Evaluation and the Complexity of Early Decision-Making. In: Public Organization Review (2017), S. 1–17
- SCHANDOCK, Manuel; BREMSER, Felix: Der Beitrag des Anerkennungsgesetzes zur Bewältigung des Fachkräftemangels. In: Berufsbildung in Wissenschaft und Praxis 41 (2012) 5, S. 11–14
- SMISMANS, Stijn: Policy Evaluation in the EU: The Challenges of Linking Ex ante and Ex Post Appraisal. In: European Journal of Risk Regulation (2015) 1, S. 6–26
- STOCKMANN, Reinhard: Was ist eine gute Evaluation? Einführung zu Funktionen und Methoden von Evaluationsverfahren. CEval Arbeitspapier 9, 2004 – URL: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-118018> (Zugriff: 23.01.2018)

Peter Jablonka, Philipp Ulmer

► Evaluation einer beruflichen Weiterbildungsmaßnahme im Kontext der deutsch-portugiesischen Berufsbildungs Kooperationen

Dargestellt wird die Evaluation einer pädagogischen Weiterbildungsmaßnahme für das betriebliche Ausbildungspersonal in Portugal, die von einer deutsch-portugiesischen Expertengruppe entwickelt und implementiert wurde. Diskutiert wird zudem, wie Nutzung, Wirksamkeit und Nachhaltigkeit von Transferaktivitäten in ein anderes Land im Bereich der beruflichen Bildung gefördert werden können.

Die Evaluation startete im Oktober 2015 und wurde Mitte 2016 abgeschlossen. Nach den Ergebnissen ist das Interesse an den Kursen sowie die Zufriedenheit mit dem Angebot sehr hoch. Tendenziell werden auch die angestrebten Wirkungen erreicht: eine Klärung der eigenen Rolle als Ausbilder/-in und eine Weiterentwicklung der berufspädagogischen Kompetenzen. Die portugiesischen Verantwortlichen haben zwar die Qualifizierungsmaßnahme und das dabei entwickelte Ausbilderprofil in das Berufsbildungssystem integriert. Auf der operativen Ebene fehlt es allerdings noch an dynamischen Strukturen.

1 Ausgangslage

Als das Projekt zur Entwicklung einer Weiterbildungsmaßnahme für das betriebliche Ausbildungspersonal (portug. *tutores*) im Rahmen der deutsch-portugiesischen Berufsbildungs Kooperation gestartet wurde, war eine Evaluierungsphase nicht vorgesehen. Wesentlicher Grund hierfür war der zeitliche Rahmen des Projekts. Die deutsch-portugiesische Zusammenarbeit hatte mit der Unterzeichnung einer Absichtserklärung am 5. November 2012 in Berlin begonnen und war zunächst für drei Jahre vereinbart worden; das Tutorenprojekt selbst lief in der zweiten Jahreshälfte 2013 an, im Anschluss an die konstituierende Sitzung der politischen Steuerungsgruppe – der sogenannten „bilateralen Arbeitsgruppe“ – im Juni 2013 in Lissabon. Die Dauer des Projekts war somit von vornherein auf zwei Jahre begrenzt. Erst als die Kooperation zwischen beiden Ländern im Juni 2015 für drei weitere Jahre ver-

längert wurde, beschloss die bilaterale Arbeitsgruppe, den entwickelten Ausbilderlehrgang evaluieren zu lassen.

Hintergrund dieser Entscheidung der bilateralen Arbeitsgruppe war in erster Linie, die Effektivität der Weiterbildungsmaßnahme zu untersuchen. Eine Bewertung der verschiedenen „Programmphasen“ des deutsch-portugiesischen Projekts – also der Wirksamkeit der „Policy“ (vgl. den Beitrag von Stockmann in diesem Band) – war somit mit dem Beschluss der bilateralen Arbeitsgruppe nicht verbunden. Eine solche Evaluation bereits zwei Jahre nach Projektbeginn wäre angesichts der komplexen Rahmenbedingungen des Projekts (wie z. B. der Mitwirkung verschiedener Institutionen und Organisationen aus beiden Ländern) sicherlich auch verfrüht gewesen.

Vier Jahre später stellt sich die Situation anders dar. Inzwischen wurde das Projekt beendet und zusätzliche Erfahrungen konnten auch im Zusammenhang mit anderen bilateralen Aktivitäten gesammelt werden. Zudem haben Deutschland und Portugal im Februar 2019 eine erneute Verlängerung ihrer Kooperation vereinbart und in diesem Rahmen die Durchführung eines Folgeprojektes im Bereich des Berufsbildungspersonals beschlossen. Ein umfassender Rückblick auf das Tutorenprojekt erscheint daher sinnvoll und wichtig, zumal mit den von Stockmann auf der Tagung vorgestellten „Schlüsselgrößen für Nachhaltigkeit“ (siehe S. 19ff. in diesem Band) nunmehr ein Analyseraster mit klar definierten Bewertungskriterien vorliegt, das eine systematische und nachvollziehbare Betrachtung erlaubt.

2 Das deutsch-portugiesische Tutorenprojekt

2.1 Berufsausbildung in Portugal

Die berufliche Bildung in Portugal ist überwiegend schulisch geprägt. Allerdings finden in den beiden wichtigsten berufsorientierten Bildungsgängen – den *Cursos Profissionais* (mit einem Anteil von 28,6 % der Jugendlichen im Sekundarbereich II) und den *Cursos de Aprendizagem* (mit einem Anteil von 6,6 % der Jugendlichen, PORDATA 2016) – praktische Phasen in Betrieben statt. Beide Bildungsgänge unterscheiden sich vor allem in Bezug auf den Anteil des betrieblichen Lernens und den Lernort Schule: In den *Cursos Profissionais* sind betriebliche Praktika von ca. zwölf Wochen (rund 10 % des Bildungsganges) vorgesehen, die entweder während der schulischen Ausbildung oder an deren Ende absolviert werden; den schulischen Unterricht erhalten die Jugendlichen vorwiegend in Berufsschulen. Demgegenüber liegt der Anteil des Lernens im Betrieb bei den *Cursos de Aprendizagem* bei ca. 40 Prozent, der schulische Teil der Ausbildung wird in Berufsbildungszentren durchgeführt. Zu den *Cursos de Aprendizagem* werden üblicherweise auch die von der Außenhandelskammer (AHK) angebotenen Berufsausbildungen gerechnet. Beide Bildungsgänge werden von verschiedenen Institutionen verwaltet: Während die *Cursos Profissionais* in den Geschäftsbereich des Bildungsministeriums gehören, liegt die Zuständigkeit für die *Cursos de Aprendi-*

zagem beim Nationalen Institut für Beschäftigung und Berufsbildung (IEFP), die wiederum dem Ministerium für Arbeit untersteht.

Wie mehrere andere europäische Staaten, vor allem im Süden des Kontinents, war auch Portugal infolge der Finanzkrise von 2008 von hoher Jugendarbeitslosigkeit betroffen. Nach Angaben von EUROSTAT waren Anfang 2014 in Portugal 35,2 Prozent der 15- bis 24-Jährigen arbeitslos. In Griechenland lag die Jugendarbeitslosenquote bei 56,2 Prozent, in Spanien bei 54,9 Prozent (EUROSTAT 2015). Um die Situation zu verbessern, leitete Portugal eine Reihe von Reformmaßnahmen ein. So wurde u. a. im Jahr 2009 der Einfluss der Sozialpartner in der beruflichen Bildung durch die Einführung von Branchenbeiräten für die berufliche Qualifikation, sogenannten „sektoralen Räten“, gestärkt. Die Räte unterstützen die Nationale Agentur für Qualifikation und berufliche Bildung (ANQEP) bei der Definition von Qualifikationen. Im gleichen Jahr wurde die Schulpflicht von neun auf zwölf Jahre erhöht. Um die Zahl der Schulabbrecher/-innen zu senken, führte das Bildungsministerium 2013 neue berufsbildende Pilotkurse (*Cursos Vocacionais*) ein. Mit den zweijährigen Kursen, die einen höheren Anteil des Lernens im Betrieb (ca. 50 %) als die beiden wichtigsten berufsorientierten Bildungsgänge vorsahen, sollte den Jugendlichen ein schnellerer Zugang zum Arbeitsmarkt ermöglicht werden (das Pilotprojekt wurde allerdings von der damaligen Linksregierung, die seit Ende 2015 im Amt war, nicht weitergeführt).

Um den Reformkurs zusätzlich zu stützen, vereinbarte das portugiesische Bildungsministerium mit dem Bundesministerium für Bildung und Forschung (BMBF) Ende 2012 eine mehrjährige Berufsbildungskoooperation. Als vorrangiges Thema wurde zu Beginn der Zusammenarbeit die Qualifizierung des Berufsbildungspersonals in den Blick genommen. Im Fokus stand dabei die Zielgruppe der betrieblichen Ausbilder und Ausbilderinnen, da sie für die Förderung arbeitsintegrierter Lernkonzepte eine Schlüsselrolle spielen.

2.2 Phasen des Tutorenprojektes (2013–2017)

Das Tutorenprojekt wurde im Herbst 2013 auf den Weg gebracht. Am Projekt waren mehrere Institutionen aus beiden Ländern beteiligt. Auf portugiesischer Seite ANQEP, IEFP sowie das Bildungsministerium, auf deutscher Seite das Bundesinstitut für Berufsbildung (BIBB), GOVET (die Zentralstelle der Bundesregierung für internationale Berufsbildungskoooperation) und das BMBF. Entwickelt wurde der Tutorenlehrgang von deutschen und portugiesischen Expertengruppen aus Praxis und Forschung. Als Bindeglied zwischen den Partnern fungierte die AHK Lissabon.

Identifizierung des Qualifizierungsbedarfs (November 2013–Juni 2014)

Die erste Phase begann mit einem intensiven Gedanken- und Informationsaustausch im Rahmen mehrerer Workshops, um die Ausbildungssysteme in beiden Ländern besser kennenzulernen. Im Frühjahr 2014 folgte eine empirische Untersuchung, um den Qualifizierungsbedarf der Tutoren/Tutorinnen zu ermitteln. Dazu wurden Interviews mit 21 Tutorinnen und

Tutoren, 17 Koordinatorinnen und Koordinatoren sowie 23 Auszubildenden in beiden berufsorientierten Bildungsgängen geführt. Die Studie, die das BIBB gemeinsam mit den portugiesischen Partnern durchgeführt hatte, ergab folgende zentrale Ergebnisse:

Obwohl beide berufsorientierten Ausbildungsmodelle grundlegende Unterschiede aufweisen, ist die Organisation der Ausbildung auf der Ebene des ausbildenden Personals in mehrfacher Hinsicht ähnlich. In den befragten Betrieben lassen sich zwei Gruppen identifizieren: Tutorinnen bzw. Tutoren in verantwortlicher, leitender Funktion und Tutorinnen bzw. Tutoren, die die Jugendlichen unmittelbar begleiten. In den Schulen bzw. Berufsbildungszentren werden die zuständigen ausbildenden Personen „Koordinatoren“ (*coordinadores*) genannt. Es handelt sich dabei um Lehrer/-innen, die neben ihrem Unterricht auch zentrale Aufgaben im Zusammenhang mit den betrieblichen Lernphasen wahrnehmen. Zum einen obliegt ihnen die Akquirierung von Ausbildungs- bzw. Praktikumsplätzen, zum anderen erstellen sie – in der Regel in Abstimmung mit den verantwortlichen Tutorinnen und Tutoren – einen sogenannten „Tätigkeitsplan“ (*plano de atividades*) für die betriebliche Phase. Grundlage des Tätigkeitsplans sind die von ANQEP definierten nationalen Ausbildungsstandards. Die Koordinatoren und Koordinatorinnen nehmen darüber hinaus eine wichtige Funktion im Bereich der Lernerfolgskontrolle ein: Ihnen berichten in regelmäßigen Abständen die Tutorinnen und Tutoren über die Leistungen, die die Jugendlichen im Betrieb gezeigt haben.

Hinsichtlich des Weiterbildungsbedarfs wünschten sich die interviewten Tutorinnen und Tutoren Unterstützung in drei Hauptbereichen:

1. Umgang mit Jugendlichen, Motivationsförderung und Bewertung von Leistungen,
2. Umsetzung curricularer Vorgaben in einen betrieblichen Ausbildungsplan, Methoden des Lernens im Prozess der Arbeit und
3. Organisation und Selbstmanagement (vgl. ULMER/MÜLLER/PIRES 2015).

Entwicklung und Erprobung des Lehrgangs (Juli 2014–Dezember 2014)

Im Juni 2014 wurden die Ergebnisse der Interviews in einem gemeinsamen Workshop mit allen Projektpartnern vorgestellt und diskutiert. Die deutsche Expertengruppe entwickelte anschließend fünf Module für die Ausbilderqualifizierung, die vor allem die berufspädagogischen Aspekte der Befragung aufgriffen:

- ▶ Modul 1: „Meine Rolle als Tutor“,
- ▶ Modul 2: „Ausbildung im Betrieb planen“,
- ▶ Modul 3: „Ausbildung im Betrieb durchführen“,
- ▶ Modul 4: „Mit Auszubildenden umgehen“ und
- ▶ Modul 5: „Auszubildende motivieren“.

Eine wichtige Grundlage für die Entwicklung der Module bildete die „Handreichung für auszubildende Fachkräfte“ (JABLONKA u. a. 2013), da die Ergebnisse der Studie in Portugal die Annahme bestätigt hatten, dass sich die pädagogischen Herausforderungen an das betriebliche Ausbildungspersonal in beiden Ländern in vielerlei Hinsicht ähneln. Die Texte und Grafiken wurden fortlaufend mit der portugiesischen Expertengruppe abgestimmt. Die Dauer des Kurses sollte 35 Stunden betragen. Die Lehrgangsdauer orientierte sich damit an dem vom Staat vorgegebenen Umfang, die Beschäftigte für ihre Weiterbildung jährlich in Anspruch nehmen können. Das Konzept sah zudem vor, dass die Schulungen der Tutorinnen und Tutoren von portugiesischen Dozentinnen und Dozenten durchgeführt werden. Von der Projektgruppe wurde daher zusätzlich ein Kurzlehrgang für portugiesische Lehrkräfte aus den berufsorientierten Bildungsgängen entwickelt. Da es sich bei dieser Gruppe um erfahrene Pädagoginnen und Pädagogen handelte, lag der Schwerpunkt des zweitägigen Kurzlehrgangs auf der Einübung von Ausbildungsmethoden zur Unterstützung des betrieblichen Lernens.

Die Erprobung des Tutorenlehrgangs begann Anfang Oktober 2014 mit einer zweitägigen Dozentenschulung. An der Schulung, die konsekutiv gedolmetscht wurde, nahmen elf schulische Lehrkräfte teil. Im direkten Anschluss daran fand der Pilotlehrgang unter der Leitung und Moderation von drei geschulten portugiesischen Lehrkräften statt. Am Lehrgang nahmen elf Tutoren und Tutorinnen teil. Während der Erprobung wurden Konzept und Lernmaterialien der Schulung zwischen den Expertengruppen aus beiden Ländern und den portugiesischen Lehrkräften fortlaufend besprochen und an die portugiesischen Rahmenbedingungen der Ausbildung angepasst.

Implementierung und Evaluation (2015–2017)

Nach der Erprobung wurde der überarbeitete Lehrgang ab dem Jahr 2015 regelmäßig von der AHK im Auftrag von IEFP landesweit angeboten. So wurden jeweils zwei Kurse im Jahr – einer im Süden des Landes (meistens in Lissabon) und einer im Norden (in Porto) durchgeführt. Diese dritte Projektphase lässt sich wiederum in zwei Perioden unterteilen: eine erste Periode mit der Implementierung des Lehrgangs, verbunden mit einer Evaluierung der Kurse von 2014, 2015 und 2016, die im nächsten Abschnitt ausführlich dargestellt wird, und eine zweite Periode, in der die portugiesischen Verantwortlichen den Tutorenlehrgang in das Berufsbildungssystem integriert und die Projektpartner sich auf ein weiteres Projekt zum Berufsbildungspersonal verständigt haben – diesmal mit dem Schwerpunkt auf die Koordinatorinnen und Koordinatoren.

3 Die Evaluation der Qualifizierungsmaßnahme: Akzeptanz und Wirkungen aufseiten der Teilnehmenden

3.1 Zweck, Aufgaben und Konzept

Gegenstand der Evaluation waren das Konzept und die Durchführung der Qualifizierung von Tutorinnen und Tutoren mit dem Fokus auf die Lehrgänge in den Jahren 2014 bis 2016. Vorrangig ging es dabei um die Verbesserung bzw. Optimierung der Maßnahme. Zu den Aufgaben der Evaluation gehörten dementsprechend:

- ▶ die Prüfung der Akzeptanz (Zufriedenheit der Teilnehmenden), Wirksamkeit und Wirkungen der Weiterbildung,
- ▶ die Ableitung von Empfehlungen zur Weiterentwicklung des Angebots und
- ▶ die Ermittlung des Bedarfs an entsprechenden Weiterbildungsangeboten (auch bei Nicht-Teilnehmenden).

Darüber hinaus sollten im Hinblick auf das weiterreichende Ziel des Projekts, nämlich die Stärkung dualer Ausbildungsstrukturen, zumindest folgende Aspekte reflektiert werden:

- ▶ die Nachhaltigkeit der Effekte und über die unmittelbare Zielgruppe der Maßnahme hinausgehende Wirkungen, z. B. die Gewinnung weiterer Kooperationspartner und die Stärkung von Netzwerkstrukturen, sowie
- ▶ die Verstetigung und der Transfer der Maßnahme, und zwar innerhalb Portugals und darüber hinaus in andere Länder mit vorwiegend (hoch)schulisch geprägtem Ausbildungssystem.

Unser Ansatz zur Wirkungsmessung orientierte sich an dem logischen Modell des von Univation¹ entwickelten „Programmbaums“ (vgl. BARTSCH u. a. 2016). Das Modell diente dazu, die Fragestellungen der Evaluierung zu präzisieren und zu systematisieren. Insbesondere bildete der Programmbaum die Grundlage für die Differenzierung und Strukturierung der Projektziele, also der angestrebten Resultate des Projekts. Zudem wurden dabei der in der Weiterbildungs-evaluation verbreiteten Vier-Ebenen-Ansatz von Kirkpatrick (Reaktion – Lernen – Verhalten – Ergebnisse; vgl. KIRKPATRICK 2006) und das CIPP-Modell von Stufflebeam (*context, input, process, product*; vgl. STUFFLEBEAM u. a. 2003) herangezogen (vgl. RÄDIKER 2011, S. 85ff.).

Diesem Ansatz folgend und entsprechend den Begriffsdefinitionen des Glossars der Evaluation von Univation (vgl. BEYWL/NIESTROJ 2009) unterscheiden wir die möglichen Wir-

1 Univation Institut für Evaluation Dr. Beywl & Associates GmbH, Köln.

kungen aufseiten der Zielgruppe (Outcome) im Hinblick auf Stabilisierungen bzw. Veränderungen

- ▶ des Wissens, der Einstellungen, der sozialen Werte oder des Könnens der Teilnehmenden (Outcome 1),
- ▶ des Verhaltens bzw. der Handlungen, also nicht nur der grundsätzlichen Bereitschaft und Fähigkeit, sondern der entsprechenden Aktivitäten (Outcome 2), und schließlich
- ▶ solche der Lebenslage bzw. des Status der Zielgruppe (Outcome 3).²

Im Rahmen der Evaluation lag der Schwerpunkt eindeutig auf der Erhebung des Outcome 1. Denn die Vermittlung von Kenntnissen und Fähigkeiten zur Bewältigung der unterschiedlichen Aufgaben als Tutorin/Tutor stand im Vordergrund der Zielsetzung der Schulungen. In begrenztem Umfang wurde auch der betriebliche Status der Befragten (Outcome 3) erfasst. Effekte auf der Verhaltensebene (Outcome 2) mittels einer standardisierten Befragung zu erfassen, schien aufgrund der unterschiedlichen Einsatzfelder und Erfahrungen der Tutorinnen und Tutoren schwierig. Daher wurden Auswirkungen der Schulung auf die Organisation und Durchführung der Ausbildung in den Betrieben exemplarisch in den qualitativen Interviews thematisiert: Was konnte von dem Erlernten in der Praxis angewendet werden?

Die Wirkungen der Qualifizierung aufseiten der Zielgruppe wurden durch folgendes methodisches Vorgehen bestimmt:

- ▶ die Selbsteinschätzung der teilnehmenden Tutoren und Tutorinnen, und zwar im Hinblick auf Veränderungen ihrer persönlichen Situation und die Bewertung ihres persönlichen Lernerfolgs;
- ▶ den Vergleich von Vorher- und Nachher-Messungen, bezogen auf ihre Rolle als Ausbilder/Ausbildlerin und auf ihre pädagogischen Fähigkeiten (wiederum auf Basis von Selbsteinschätzungen), und
- ▶ durch Triangulation, d. h. in diesem Falle ergänzende qualitative Befragungen von Teilnehmenden und ebenfalls qualitative Interviews mit den Dozentinnen zur Validierung, Vertiefung und Veranschaulichung der Ergebnisse der standardisierten Befragungen.

Ursprünglich war zudem geplant, im Sinne eines quasiexperimentellen Designs (ohne Randomisierung) die Befragung einer Kontrollgruppe in die Untersuchung miteinzubeziehen.

2 Wirkungen, die über die unmittelbare Zielgruppe der Qualifizierungsmaßnahme hinausgehen und übergreifende Organisationen/Institutionen wie Ausbildungsbetriebe, berufliche Schulen bzw. Ausbildungszentren oder gar das System der beruflichen Ausbildung betreffen (im Sinne eines *Impacts* des Projekts), werden im Zusammenhang mit den Aspekten „Verstetigung“ (Abschnitt 4.1) und „Transfer“ (Abschnitt 4.2) thematisiert.

Die Beteiligung an dieser Erhebung fiel allerdings so gering aus (siehe Abschnitt 3.2), dass die entsprechenden Daten nicht berücksichtigt werden konnten.

3.2 Erhebungen und Stichproben

Die Evaluation war summativ angelegt und verknüpfte den Einsatz von quantitativen und qualitativen Methoden. Durchgeführt wurden:

- ▶ standardisierte Online-Erhebungen unter Teilnehmenden an der Weiterbildung, und zwar kurz vor deren Beginn und ca. sechs Monate nach deren Abschluss, sowie
- ▶ leitfadengestützte Interviews mit Dozentinnen und Teilnehmenden ca. sechs Monate nach der Weiterbildung.

Dabei widmete sich eine erste Untersuchung den beiden Kursen, die Ende 2015 durchgeführt wurden, und ergänzend der Pilotierung des Lehrgangs von 2014. Die erste Befragungswelle wurde unmittelbar vor Kursbeginn im November 2015 gestartet, die zweite folgte ca. ein halbes Jahr später. Über die Ergebnisse wurde zunächst im August 2016 berichtet. Um die Erkenntnisse der Untersuchung auf eine breitere Basis zu stellen, wurden zusätzlich die bis Ende 2016 durchgeführten Lehrgänge in die Evaluierung miteinbezogen.

Angeschrieben und um eine Beteiligung an der Befragung gebeten wurden im Rahmen der ersten Welle (vor der Schulung) alle 60 Teilnehmerinnen und Teilnehmer an den vier Lehrgängen der Jahre 2015 und 2016. In die zweite Welle (nach der Schulung) wurden zusätzlich elf Teilnehmer/-innen des Pilot-Lehrgangs von 2014 einbezogen. Beteiligt haben sich an den Befragungen bei der ersten Welle 42 und bei der zweiten 29 Tutorinnen und Tutoren. Die entsprechenden Rücklaufquoten lagen damit bei 75 Prozent für die erste und 41 Prozent für die zweite Welle. Hinzuweisen ist an dieser Stelle darauf, dass – im Hinblick auf die Einschätzung der Wirksamkeit der Schulung – der Fokus auf der Gruppe der Tutoren und Tutorinnen lag, die sich an beiden Befragungswellen, also vor und nach den Lehrgängen, beteiligt haben. Der Umfang dieser Teilstichprobe lag allerdings nur bei 22 Personen. Einzelheiten zum Einsatz und Rücklauf aller Erhebungen werden in der Tabelle dokumentiert

Die qualitativen Erhebungen wurden nur im Rahmen der ersten Studie durchgeführt. Befragt werden konnten zehn Teilnehmer/-innen und zwei Dozentinnen.

Wie erwähnt, war ursprünglich auch eine Befragung von Tutoren und Tutorinnen geplant, die nicht an den Qualifizierungsmaßnahmen teilgenommen hatten. Dazu konnten mit Unterstützung der AHK insgesamt 370 Personen, die der Kammer als Tutor/Tutorin oder als für die Ausbildung in ihren Unternehmen Zuständige bekannt waren, kontaktiert werden. Der Rücklauf fiel allerdings mit 14 auswertbaren Fällen bzw. vier Prozent so gering aus, dass auf eine Auswertung dieser Befragungsdaten verzichtet wurde.

Tabelle 1: Einsatz von Adressen und Rücklauf bei der Teilnehmendenbefragung

		Eingesetzte Adressen	Rücklauf	
		Anzahl	Anzahl	%
Welle 1		60	42	75
Welle 2, insgesamt		71	29	41
davon	Teilnehmende in 2016	24	8	33
	Teilnehmende in 2015	36	17	47
	Teilnehmende in 2014	11	4	36
	Teilnehmende an Welle 1 und 2	60	22	37
	Teilnehmende nur an Welle 2	60	3	5

3.3 Vielfältige und herausfordernde Aufgaben

Bei den Teilnehmern und Teilnehmerinnen der Schulungen handelte es sich in Bezug auf Erfahrungen, Qualifikation und berufliche Stellung wie auch die Branche bzw. das Berufsfeld und die konkreten betrieblichen Aufgaben als Tutor bzw. Tutorin um eine sehr heterogene Gruppe. Dieser Vielfalt gerecht zu werden, stellte besondere Anforderungen an die Dozentinnen, war jedoch dem Erfahrungsaustausch und dem Voneinander-Lernen in der Gruppe sehr förderlich.

Auf den ersten Blick hatten nur kleinere Gruppen größere Schwierigkeiten bei der Bewältigung ihrer einzelnen Aufgaben als Tutor/-in. Allerdings sind es jeweils durchaus relevante Gruppen, die einräumen, zumindest teil- und zeitweise nicht allen ihren Anforderungen gerecht geworden zu sein. Dabei stehen zum einen die innerbetriebliche Organisation und die Koordinierung von betrieblicher und schulischer Ausbildung im Vordergrund; zum anderen geht es vor allem um erzieherische und disziplinarische Aufgaben (auch Motivation der Auszubildenden).

3.4 Hohe Motivation und Zufriedenheit

Das Interesse an dem Lehrgang und allen einzelnen angebotenen Themen war ebenso wie die Motivation, etwas zu lernen und die Ausbildung im Betrieb zu verbessern, bei allen Teilnehmenden hoch. Soweit dazu Informationen vorliegen, haben bis auf wenige Ausnahmen alle den Kurs bis zum Ende besucht und mit Erfolg abgeschlossen – auch wenn sie an einzelnen Stunden (vor allem aufgrund beruflicher Verpflichtungen) nicht teilnehmen konnten.

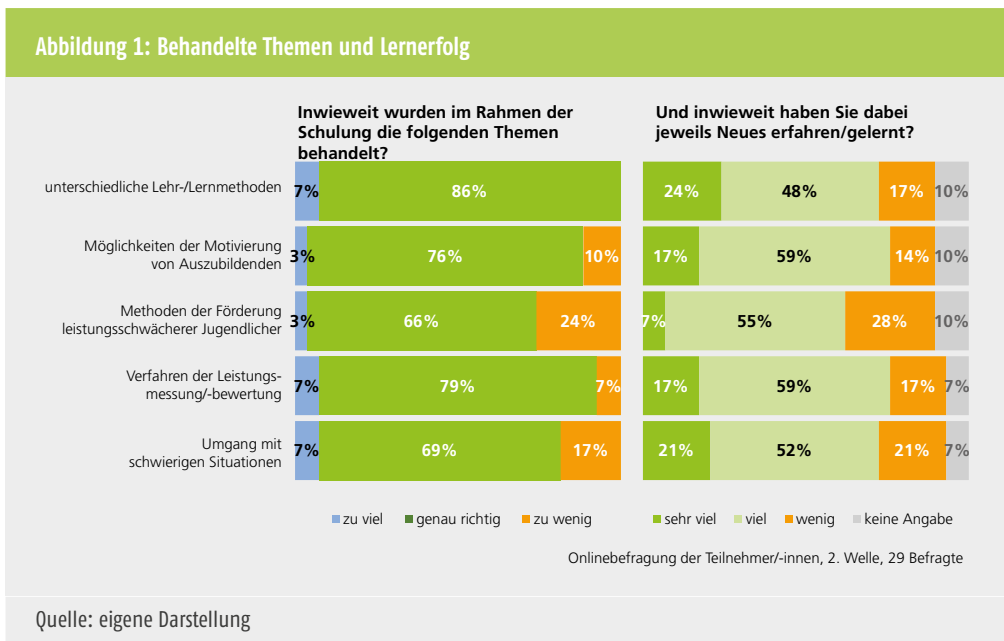
Fast alle Teilnehmenden (ca. 90 %) waren mit dem Kurs insgesamt zufrieden oder sehr zufrieden. Ausnahmslos wurden die Dozentinnen positiv beurteilt – und zwar sowohl im Hinblick auf ihr fachliches Wissen wie auch die Vermittlung der Kursinhalte und nicht zuletzt die Einbeziehung und Aktivierung der Teilnehmenden. Weit überwiegend wurde auch die Intensität, mit der die einzelnen Inhalte behandelt wurden, als angemessen und

der Informationsgehalt der behandelten Themen als hoch bezeichnet. Die Dauer des Kurses (35 Stunden) hielten die allermeisten (86 %) für gerade richtig. Auch an der Verteilung der Lehrgangzeiten auf vier Wochen mit jeweils zwei Vormittagen pro Woche gab es nichts auszusetzen. Weitere Aspekte des Kurses, wie Verständlichkeit, Lehrgangsmaterialien, räumliche Ausstattung oder Praxisnähe bekommen ebenfalls überwiegend gute Noten.

3.5 Wirkungen der Qualifizierung: Lernerfolg und Statussicherung

Wie erwähnt (siehe Abschnitt 3.1) wurden die Wirkungen der Qualifizierung aufseiten der Tutorinnen und Tutoren durch deren Selbsteinschätzung des Lernerfolgs, den Vergleich von Vorher- und Nachher-Messungen sowie ergänzende qualitative Befragungen erfasst. Im Folgenden sollen dazu exemplarisch einzelne Ergebnisse dargestellt werden, zunächst auf Basis der Selbsteinschätzung der Befragten.

Die große Mehrheit der 29 Teilnehmer/-innen der zweiten Befragungswelle bewertet ihren persönlichen Lernerfolg positiv: Dies gilt im Hinblick auf den gesamten Kurs – 89 Prozent von ihnen geben an, sehr viel (41 %) oder viel (48 %) gelernt zu haben – und auch bezogen auf einzelne Themen der Schulung (siehe Abb. 1).

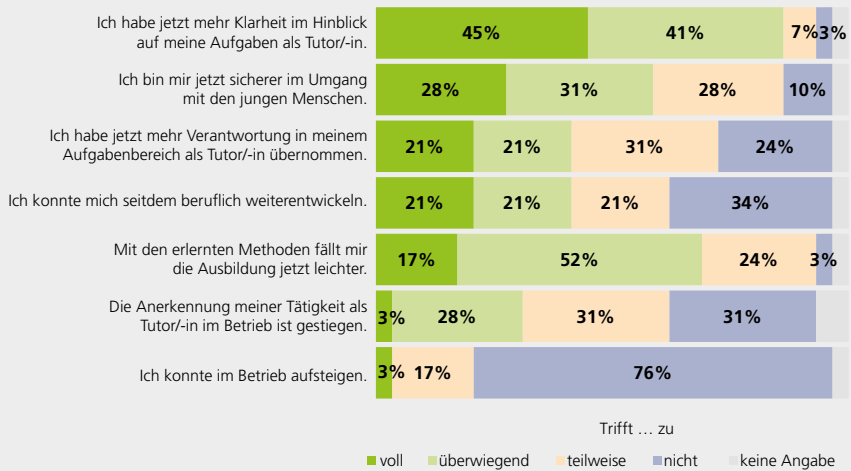


Darüber hinaus hatte die Schulung auch Auswirkungen auf die Selbstwahrnehmung der eigenen Rolle und (begrenzt) die berufliche Situation (siehe Abb. 2). Vor allem haben die Befragten nach Abschluss der Schulung mehr Klarheit im Hinblick auf ihre Aufgaben als Tu-

tor/Tutorin und mehr Sicherheit im Umgang mit jungen Menschen gewonnen. Auch meinen viele, dass ihnen aufgrund der erlernten Methoden die Ausbildung jetzt leichter fiele. Bei Einzelnen hat sich auch ihre Position bzw. ihr Status im Betrieb verändert: Sie haben mehr Verantwortung übernommen oder die Anerkennung ihrer Tätigkeit als Tutor/-in ist gestiegen.

Abbildung 2: Veränderungen nach der Schulung

Was hat sich seit Ihrer Teilnahme an der Schulung verändert?



Onlinebefragung der Teilnehmer/-innen, 2. Welle, 29 Befragte

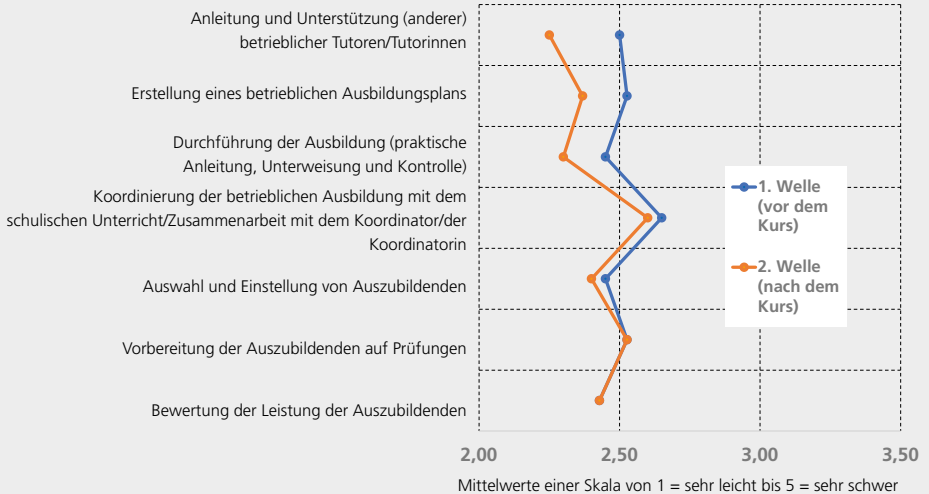
Quelle: eigene Darstellung

Eine ganze Reihe von Fragen wurde den Teilnehmern und Teilnehmerinnen (in gleicher Formulierung) sowohl kurz vor der Schulung wie auch im Abstand von ca. sechs Monaten danach gestellt. Dargestellt werden soll im Folgenden, wie sich die Fähigkeit der Tutoren und Tutorinnen, einzelne Aufgaben aus ihrem Ausbildungsalltag zu bewältigen, im Vergleich beider Befragungswellen verändert hat – dabei werden entsprechende Veränderungen als Wirkungen der Qualifizierung interpretiert.

Die abgefragten Ausbildungsaufgaben wurden in zwei Blöcken präsentiert. Im ersten Block (siehe Abb. 3) standen Anforderungen im Vordergrund, die eher sporadisch anfallen und zum Teil organisatorischen/koordinierenden Charakter haben (z. B. Einstellung von Auszubildenden oder Planung der Ausbildung).

Abbildung 3: Vorrangig sporadisch anfallende Ausbildungsaufgaben

Und wie leicht bzw. schwer fällt Ihnen die Bewältigung dieser Aufgaben?



Onlinebefragung der Teilnehmer/-innen, 1. und 2. Welle, 19 bis 21 Befragte

Quelle: eigene Darstellung

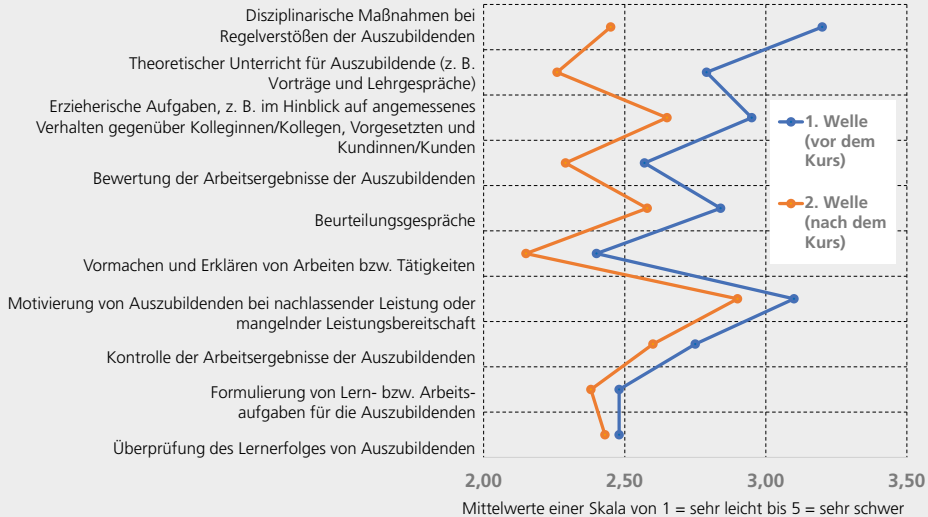
Anschließend wurden die Befragten gebeten einzuschätzen, wie leicht bzw. schwer ihnen die Bewältigung eher häufig und regelmäßig anfallender Aufgaben falle (siehe Abb. 4).

Die Unterschiede hinsichtlich der Selbsteinschätzung der Fähigkeiten zu den beiden Befragungszeitpunkten sind zwar zum Teil gering (und weitgehend statistisch nicht signifikant), sie weisen aber (nahezu) durchgängig in die erwartete Richtung: Den Tutoren und Tutorinnen fällt die Bewältigung ihrer Ausbildungsaufgaben im Abstand von ca. einem halben Jahr nach Abschluss ihres Lehrgangs leichter als zuvor. Dies gilt insbesondere für Anforderungen aus dem Arbeitsalltag, tendenziell auch für einzelne planerische bzw. organisatorische Aufgaben.

Ein Vergleich der Skalenmittelwerte lässt allerdings nicht ohne Weiteres Rückschlüsse auf individuelle Veränderungen zu. Deshalb wurde im Rahmen einer weiteren Auswertung der Daten untersucht, inwieweit es, bezogen auf die jeweils einzelnen Lehrgangsteilnehmer/-innen, eine Entwicklung gegeben hat. Abbildung 5 weist aus, bei wie vielen Tutoren und Tutorinnen es beim Vergleich beider Befragungszeitpunkte eine „Verbesserung“ bzw. eine „Verschlechterung“ der Selbsteinschätzung ihrer Fähigkeiten gegeben hat.

Abbildung 4: Regelmäßig anfallende Ausbildungsaufgaben

Und wie leicht bzw. schwer fällt Ihnen die Bewältigung dieser Aufgaben?



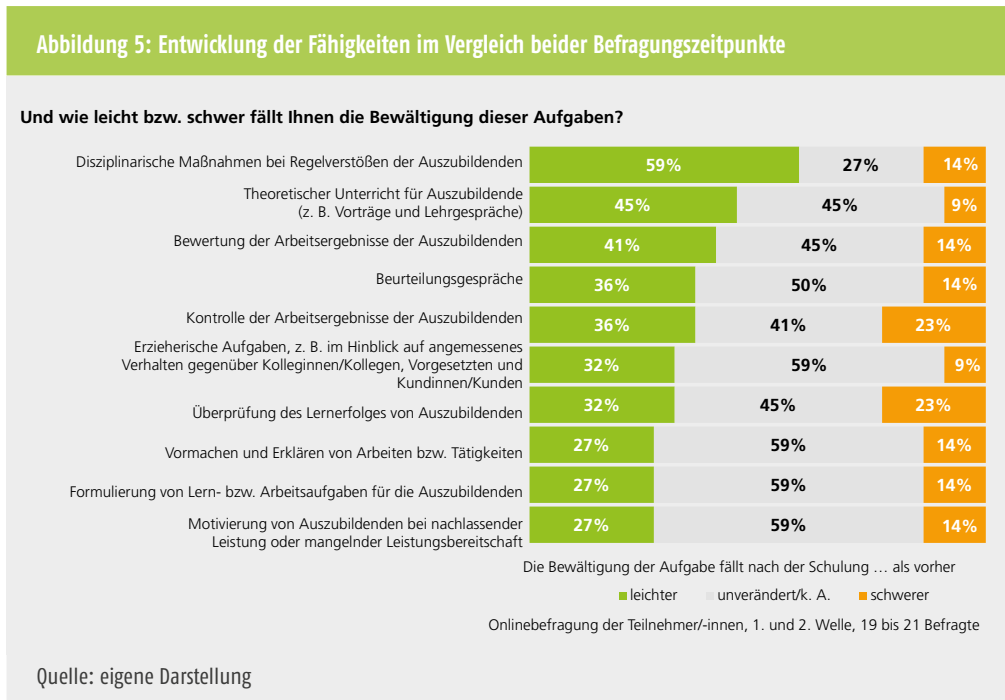
Onlinebefragung der Teilnehmer/-innen, 1. und 2. Welle, 19 bis 21 Befragte

Quelle: eigene Darstellung

Zunächst einmal fällt auf, dass – entsprechend der Tendenz der Mittelwertsunterschiede – jeweils etwa ein Viertel bis (gut) die Hälfte der Tutoren/Tutorinnen sich „verbessert“ hat. Allerdings gibt es auch einzelne Befragte, die ihre Fähigkeiten nach dem Kurs geringer einschätzen als zuvor. Dies könnte als Hinweis auf eine mangelnde Zuverlässigkeit des Messinstruments gedeutet werden (und würde in gleicher Weise für Veränderungen in entgegengesetzter Richtung gelten). Aber die entsprechenden Ergebnisse insbesondere im Hinblick auf die Kontrolle bzw. Überprüfung des Lernerfolgs bzw. der Arbeitsergebnisse der Auszubildenden lassen auch eine andere Interpretation zu: Plausibel erscheint, dass den Lehrgangsteilnehmern und -teilnehmerinnen erst durch die Behandlung dieser Themen im Kurs die Problematik dieser Anforderungen und infolgedessen ihre bisherigen Schwächen bei deren Bewältigung bewusst geworden sind. Dies gilt umso mehr, als sowohl bei der Konzeption wie auch in der Wahrnehmung der Teilnehmenden (dies zeigten die qualitativen Interviews) der Erfahrungsaustausch, das Feedback vonseiten der Dozentinnen und der Kollegen/Kolleginnen und die Reflexion des eigenen Verhaltens einen breiten Raum innerhalb der Qualifizierung einnahm.

Die Ergebnisse der Leitfadenterviews zeigen schließlich, dass die Tutoren und Tutorinnen zum Teil sehr konkrete Hilfestellungen und Arbeitsunterlagen aus dem Kurs mitge-

nommen und in ihrem Arbeitsalltag genutzt haben. Genannt werden z. B. Arbeitshilfen zur Systematisierung und Organisation der Ausbildung oder von Praktika (insbesondere zur Erstellung von Ausbildungsplänen), zur Durchführung regelmäßiger Feedbackrunden oder Beurteilungsgespräche sowie zur Abstimmung von Ausbildungsinhalten und Methoden auf die spezifischen Anforderungen unterschiedlicher Zielgruppen.



Zudem wurde deutlich, dass die Weiterbildungsmaßnahme tatsächlich dazu geführt hat, das eigene Handeln zu überdenken und sich neu zu orientieren. So merkt eine Tutorin an, dass sie jetzt „einen anderen Blick auf ihre Tätigkeit als Tutorin“ habe. Sie wisse nun genau, was sie tue, und habe gelernt, dass sie den Auszubildenden zuvor zu wenig erklärt hatte. Sie habe zu viel vorausgesetzt, habe sich nicht so vor Augen geführt, dass die Auszubildenden vieles noch nicht wussten. So sei sie bemüht, bei ihren derzeitigen Auszubildenden umzusetzen, was sie gelernt habe, und Fehler zu vermeiden, die sie vorher gemacht habe.

Der Lehrgang hat demnach dazu beigetragen, dass die teilnehmenden Tutoren und Tutorinnen ihre bisherige Praxis kritisch hinterfragt haben, die betriebliche Ausbildung neu und besser organisiert haben, einzelne erlernte Methoden und Techniken anwenden sowie Kollegen und Kolleginnen bei der Ausbildung unterstützen konnten.

3.6 Schlussfolgerungen und Empfehlungen zur Fortführung der Qualifizierungsmaßnahme

Das Konzept der Schulung hat sich grundsätzlich inhaltlich wie im Hinblick auf die Struktur und Organisation des Kurses bewährt. Es erscheint geeignet, die Qualität der Ausbildung in den Betrieben zu verbessern. Die Schulung sollte daher in dieser Form in Zukunft weiter angeboten werden.

Die Erwartungen und Ausbildungsbedarfe der Tutoren/Tutorinnen werden durch die Themenbereiche der fünf Module prinzipiell voll abgedeckt. Zu berücksichtigen ist allerdings, dass Tutoren und Tutorinnen häufig nicht nur für Auszubildende, sondern auch für die Anleitung von Praktikanten/Praktikantinnen (z. B. Umschüler/-innen und Studierende) zuständig sind. Zu prüfen ist, inwieweit hier eine Ergänzung der Schulungsinhalte sinnvoll ist und ggf. das Thema der Kooperation der Lernorte zu akzentuieren ist.

Die Durchführung der Schulung stellt allerdings hohe Anforderungen an die Dozenten und Dozentinnen: Sie müssen über entsprechende Erfahrungen und Kompetenzen in der beruflichen Ausbildung wie auch in der Erwachsenenpädagogik verfügen. Insbesondere müssen sie in der Lage sein, mit der Heterogenität der Teilnehmenden, vor allem ihrer unterschiedlichen Herkunft (Branchen und Unternehmensgrößen) und ihren unterschiedlichen Aufgabenstellungen, angemessen umgehen zu können. Eine entsprechende Fortbildung geeigneter Lehrkräfte erscheint daher geboten.

Es liegen keine Informationen dazu vor, inwieweit der Lehrgang vom zeitlichen Umfang und der zeitlichen Lage her den Möglichkeiten und Anforderungen einer möglichst großen Zahl von Tutoren und Tutorinnen entspricht. Zumindest stieß er bisher auf erheblichen Zuspruch der Zielgruppe. Zu prüfen bleibt jedoch, ob auch andere organisatorische Angebote (Abend- bzw. Ganztagskurse oder eine Aufteilung in einzelne Module) auf eine ausreichende Nachfrage stoßen würden.

In den qualitativen Interviews hatten die Teilnehmer und Teilnehmerinnen der Schulung Interesse an einem weiterführenden, vertiefenden Kurs gezeigt. Solch ein Lehrgang sollte die praktischen Elemente (z. B. Simulationen, Rollen- und Planspiele) verstärken und dabei insbesondere die Erfahrungen und Interessen der Teilnehmenden einbeziehen. Von einzelnen Befragten wird die Entwicklung eines entsprechenden didaktisch-methodischen Konzepts empfohlen.

4 Wirkungen des Projekts, die über die Zielgruppe hinausreichen

4.1 Verstetigung und Fortentwicklung des Qualifizierungsangebots

Die Ergebnisse der Evaluierung lassen den Schluss zu, dass eine breite Implementierung der Tutorenschulung den Grundstein dafür legen kann, eine stärkere Praxisorientierung in der Ausbildung zu fördern. Im Jahr 2017 wurden daher weitere Schritte zur Verstetigung der Qualifizierungsmaßnahme für Tutoren eingeleitet:

- ▶ Die Schulung wurde ins portugiesische Berufsausbildungssystem integriert, nachdem der Aufbau noch einmal leicht verändert wurde. In ihrer endgültigen Form besteht sie nunmehr aus vier Modulen à zehn Stunden: (1) „Rolle und Aufgaben des Tutors“, (2) „Instrumente zur Unterstützung der pädagogischen Praxis“, (3) „individueller Ausbildungsplan“ und (4) „Motivation in der betrieblichen Ausbildung“ (<https://www.iefp.pt/referenciais-formacao/-/sgd/get-file/690358/690358>).
- ▶ Neben der Schulung veröffentlichte IEFP auf seiner Homepage auch ein Anforderungs- und Kompetenzprofil für betriebliche Tutoren.
- ▶ Darüber hinaus wurde auf der Grundlage des Kurzlehrgangs für Lehrkräfte vom Oktober 2014 (siehe Abschnitt 2.2 zur zweiten Projektphase) eine Dozentenschulung von der AHK Lissabon entwickelt. Die Schulung war von IEFP initiiert worden (<https://www.iefp.pt/referenciais-formacao/-/sgd/get-file/690424/690424>).

Bis heute finden regelmäßig jedes Jahr mehrere Tutorenlehrgänge statt. Neben der AHK Lissabon, die die meisten Kurse durchführt, gibt es auch portugiesische Anbieter.

Um die betriebliche Ausbildung in Portugal weiter zu fördern, beschlossen die Partner, Unterstützungsmaßnahmen auch für Koordinatorinnen und Koordinatoren, die zweite wichtige Gruppe des Berufsbildungspersonals im portugiesischen Berufsausbildungssystem, zu ergreifen. In Anlehnung an das Tutorenprojekt sollten auch für Koordinatorinnen und Koordinatoren ein Anforderungs- und Kompetenzprofil erstellt sowie eine Fortbildungsmaßnahme entwickelt werden. Das neue Projekt, das im Dezember 2017 startete, hat jedoch auf der operative Ebene eine andere Struktur. Im Gegensatz zur Weiterbildungsmaßnahme für Tutoren und Tutorinnen ist die AHK Lissabon bei der Projektgestaltung nicht eingebunden. Deren Rolle als „Antreiberin“ für die Umsetzung verabredeter Schritte – die AHK hatte hierfür im Tutorenprojekt eigens mehrere zweisprachige Mitarbeiterinnen eingesetzt – sollten die portugiesischen Institutionen übernehmen. Dem konnten die Institutionen u. a. aufgrund personeller Wechsel und begrenzter Ressourcen bis jetzt allerdings nur bedingt gerecht werden, was auch zu mehreren Verzögerungen in der Projektdurchführung führte.³

4.2 Voraussetzungen für einen Transfer dualer Strukturelemente

Die bisherigen Erfahrungen in der bilateralen Kooperation zeigen, dass der Verlauf von internationalen Berufsbildungsprojekten von mehreren Faktoren abhängt. Die Frage, wann ein Projekt als Erfolg oder als Misserfolg eingestuft wird und welche Kriterien hierfür herangezogen werden sollen, dürfte je nach Perspektive und Interessenlage unterschiedlich beantwortet werden. Mit den von Stockmann vorgestellten „Schlüsselgrößen für Nachhaltigkeit“, die mithilfe einer Metaanalyse von 83 internationalen Projekten und Programmen der

3 Dies war der Stand bis zum Zeitpunkt der Fertigstellung des Beitrags (Sommer 2020). Inzwischen haben die portugiesischen Institutionen reagiert und die AHK in die Gestaltung des Projekts wieder eingebunden.

Berufsbildungszusammenarbeit herausgearbeitet wurden (siehe S. 21 in diesem Band), liegt nun aber ein umfassender, klar definierter Kriterienkatalog vor, mit dem eine nachvollziehbare Bewertung von bilateralen Berufsbildungsprojekten vorgenommen werden kann.

Betrachtet man nun beide Projekte zum Berufsbildungspersonal im Lichte der vier „Schlüsselgrößen“ und ihren Unterpunkten lässt sich Folgendes feststellen:

Das Tutorenprojekt kann insgesamt als erfolgreich eingestuft werden. Es wurde allen vier Schlüsselgrößen „Systemkompatibilität“, „Flexible Steuerung“, „Personal“ und „Ownership“ weitestgehend gerecht. Durch die Workshops zu Beginn der Zusammenarbeit, die empirische Untersuchung über den Qualifikationsbedarf von Tutoren und Tutorinnen sowie die enge und kontinuierliche Abstimmung mit den portugiesischen Bildungsfachleuten konnte die Schulung an das Berufsbildungssystem in Portugal gut angepasst werden (vgl. Abschnitt 2.2). Immer wieder wurden auch die einzelnen Etappen des Projekts sowie die inhaltlichen Aspekte der Schulung mit den portugiesischen Institutionen rückgekoppelt. Dies betraf sogar die Übersetzung einzelner Begriffe. Dank des regelmäßigen und sehr flexiblen Austauschs innerhalb der Steuerungsgruppe (BIBB und AHK) konnten organisatorische und inhaltliche Fragen zwischen den Partnern schnell und effektiv geklärt werden. Eine entscheidende Voraussetzung hierfür war auch der Umstand, dass die AHK teilweise mehrere Mitarbeiter und Mitarbeiterinnen – oft zweisprachig – in das Projekt eingebunden hatte. Durch diese Maßnahme, die vom BMBF finanziell unterstützt worden war, konnte der dritte Schlüsselfaktor „Personal“ erfüllt werden (was sich auch auf die „flexible Steuerung“ sehr positiv ausgewirkt hatte). Dies hatte darüber hinaus einen erheblichen Einfluss auf die vierte Schlüsselgröße „Ownership“. Da die AHK selbst drei Ausbildungszentren unterhält, hatte sie ein starkes Eigeninteresse an der Entwicklung einer Qualifikation für das Ausbildungspersonal. Zu diesem vierten Erfolgsfaktor zählt Stockmann auch den „Veränderungswillen“ der Verantwortlichen im Partnerland. Die Schritte, die IEFB vor allem nach der Entwicklung der Tutorenschulung auf den Weg gebracht hat (vgl. Abschnitt 4.1), zeigen, dass auch dieser Faktor wirksam war.

Das Koordinatorenprojekt lässt naturgemäß noch keine abschließende Bewertung zu. Die Laufzeit des Projektes erstreckt sich bis Ende 2021 bzw. Anfang 2022. Es zeigt sich aber schon jetzt, dass die fehlende Einbindung der AHK in die Gestaltung des Projektes sich auf alle vier Schlüsselgrößen nachteilig auswirkt. Dies betrifft vor allem die Steuerung des Projektes, die personellen Ressourcen auf portugiesischer Seite und den Faktor „Ownership“, die bei diesem zweiten Projekt schwächer ausgeprägt sind.⁴

Der Rückblick auf beide Projekte mithilfe der von Stockmann entwickelten Kriterien macht zugleich deutlich, welche Rahmenbedingungen gegeben sein sollten, um die Chancen für einen Transfer dualer Strukturelemente im Rahmen von bilateralen Berufsbildungsoperationen zu verbessern. Aus unserer Sicht sind die Schlüsselgrößen und ihre Unterpunkte zentrale Voraussetzungen dafür.

4 Dies war der Stand bis zum Zeitpunkt der Fertigstellung des Beitrags (Sommer 2020). Inzwischen haben die portugiesischen Institutionen reagiert und die AHK in die Gestaltung des Projekts wieder eingebunden.

5 Fazit

Die Evaluation der Qualifizierungsmaßnahme hat zwar durchaus relevante Erkenntnisse zur Akzeptanz und (begrenzt) zur Wirkung der Schulungen erbracht, sie ist allerdings im Hinblick auf folgende Aspekte auch kritisch einzuschätzen:

- ▶ Die Stichprobe war, insbesondere was die Teilnahme an beiden Befragungswellen angeht, so klein, dass insbesondere die Resultate zur Kompetenzentwicklung der Teilnehmenden nur wenig belastbar sind. Problematisch in diesem Zusammenhang war, dass aufseiten des Auftraggebers und des Evaluators vorab keine Informationen über den Umfang der zu evaluierenden Maßnahme (Anzahl der geplanten Kurse und zu erwartenden Teilnehmer/-innen) vorlagen und auch kein Einfluss auf die Planung genommen werden konnte.
- ▶ Die Aussagekraft der Studie ist zudem dadurch begrenzt, dass entgegen der ursprünglichen Absicht, keine Tutorinnen und Tutoren einbezogen werden konnten, die nicht an einer Schulung teilgenommen haben. Dadurch konnte nicht eingeschätzt werden, ob und inwieweit sich die Teilnehmenden von der Zielgruppe der Maßnahme insgesamt unterscheiden (z. B. hinsichtlich ihrer betrieblichen Stellung und ihrer Aufgaben im Rahmen der betrieblichen Ausbildung) und inwieweit generell Interesse an einer entsprechenden Qualifizierung besteht.
- ▶ Aufgrund der begrenzten Laufzeit der Evaluation und der Konzentration auf die Durchführung der Schulungen und deren Wirkung auf die Zielgruppe konnten zudem weiterreichende Effekte, z. B. auf der Ebene der Ausbildungsorganisation in den Betrieben oder im Hinblick auf die Kooperation mit Ausbildungszentren bzw. beruflichen Schulen, nicht untersucht werden.

Das heißt, letztlich waren auf dieser Basis auch keine belastbaren Aussagen dazu möglich, inwieweit die Weiterbildungsmaßnahme tatsächlich zur Festigung oder gar zum Ausbau der betrieblichen Ausbildung in Portugal beitragen konnte.

Allerdings haben die Ausführungen zu den Themen „Verstetigung“ und „Transfer“ gezeigt, dass im Rahmen der Fortführung und Weiterentwicklung des Projekts sowie der Beobachtung und Reflexion der Resonanz aufseiten der Kooperationspartner, also außerhalb und ergänzend zur systematischen Evaluation der Schulungen, durchaus entsprechende Erkenntnisse gesammelt werden konnten: Wirkungen, die nicht nur bei einer begrenzten Zahl von Akteuren der beruflichen Bildung greifen, sondern auf eine Stärkung bzw. den Ausbau von zumindest einzelnen Elementen betrieblicher Ausbildung gerichtet sind, sind nur mit langfristiger Perspektive zu erwarten, und sie werden in starkem Maße von den Rahmenbedingungen aufseiten der Kooperationspartner bestimmt.

Literatur

- BARTSCH, Samera; BEYWL, Wolfgang; NIESTROJ, Melanie: Der Programmbaum als Evaluationsinstrument. In: GIEL, Susanne; KLOCKGETHER, Katharina; MÄDER, Susanne (Hrsg.): Evaluationspraxis. Professionalisierung – Ansätze – Methoden. Münster, New York 2016, S. 89–112
- BEYWL, Wolfgang; NIESTROJ, Melanie: Das A-B-C der wirkungsorientierten Evaluation. Glossar Deutsch / Englisch der wirkungsorientierten Evaluation. 2. Aufl. Köln 2009
- EUROSTAT (Hrsg.): Arbeitslosenquote des Euroraums bei 11,2 %. EU28 bei 9,8 %. pressemitteilung euroindikatoren 36/2015 am 2. März 2015, S. 5 – URL: <https://ec.europa.eu/eurostat/documents/2995521/6664120/3-02032015-AP-DE.pdf/0e4e0fde-e78c-40b2-bbdb-fe617328820d> (Zugriff: 09.09.2019)
- JABLONKA, Peter; MARTIN, Sigrid; ULMER, Philipp: Handreichung für ausbildende Fachkräfte. Bonn 2013
- PORDATA (Hrsg.): Base de Dados Portugal Contemporâneo – URL: <https://www.pordata.pt/Portugal/Alunos+matriculados+no+ensino+secundário+total+e+por+modalidade+de+ensino-1042> (Zugriff: 05.10.2017)
- RÄDIKER, Stefan: Die Evaluation von Weiterbildungsprozessen in der Praxis: Status quo, Herausforderungen, Kompetenzanforderungen. Marburg/Lahn 2012
- STOCKMANN, Reinhard: Systematic Evaluation Analysis (SEA) als Instrument der Wirkungsevaluation. Vortrag auf dem AGBFN-Forum „Evaluation und Wirkungsforschung in der Berufsbildung“ in Bonn am 29. November 2018 – URL: https://www.agbfn.de/dokumente/pdf/AGBFN_Wirkungsforschung_Praesentation_Stockmann.pdf (Zugriff: 25.02.2019)
- STUFFLEBEAM, Daniel L.; MCKEE, Harold; MCKEE, Beulah: The CIPP Model for Evaluation. An update. A review of the model's development. Presented at the Annual Conference of the Oregon Program Evaluators Network (OPEN). Portland, Oregon 2003
- ULMER, Philipp; MÜLLER, Hans-Joachim; PIRES, Fatima: Berufspädagogische Weiterbildung betrieblicher Tutorinnen und Tutoren. Ein deutsch-portugiesisches Pilotprojekt. In: Berufsbildung in Wissenschaft und Praxis 44 (2015) 4, S. 34–37

Birte Komosin, Henning Kruse

► **Welchen Beitrag kann ein regionales Monitoring- und Berichtssystem für die Wirkungsforschung in der Berufsbildung leisten?**

In Berlin kommt ein digitales Monitoring- und Berichtssystem zur Anwendung, über das Ergebnisse und Wirkungen öffentlich geförderter Beratungs- und Coachingleistungen in der beruflichen Weiterbildung und der Arbeitsförderung kontinuierlich erfasst werden. Über dieses Monitoringverfahren werden Erkenntnisse zur Zielerreichung der entsprechenden Förderinstrumente generiert. Außerdem können erste Hinweise zur Wirksamkeit der Angebote gewonnen werden. Durch die konsequente Ausrichtung des Verfahrens auf ein digitales System ist es in hohem Maße skalierbar und effizient. Der Beitrag skizziert diesen Ansatz und erörtert, wie ein solches Monitoring- und Berichtssystem für weitergehende Evaluierungen und Wirkungsuntersuchungen nutzbar gemacht werden kann.

1 Einführung

Beratungs- und Coachingleistungen im Bereich der beruflichen Weiterbildung und der Arbeitsförderung sind ein wichtiger Baustein einer aktiven Arbeitsmarkt- und Berufsbildungspolitik. Ziel dieser Angebote ist es, die Menschen dabei zu unterstützen, sich beruflich weiterzuentwickeln, und ihnen bei der Integration in Arbeit zu helfen. Zu solchen Angeboten gehören beispielsweise Bildungs- und Berufsberatungen, Coachingmaßnahmen zur Arbeitsmarktintegration und zum Abbau von Vermittlungshemmnissen für Langzeitarbeitslose, die Anerkennungsberatung für Menschen mit Flucht- und Migrationshintergrund. Für die Legitimierung und Finanzierung sowie die Ausrichtung und Steuerung der Maßnahmen benötigen die Politik ebenso wie die Beratungseinrichtungen Informationen zu ihren Ergebnissen und zu ihrer Wirksamkeit. Gerade Untersuchungen zur Messung von Wirksamkeit in Form von Evaluierungen und Wirkungsanalysen stellen Politik und Wissenschaft jedoch vor große Herausforderungen. Der hohe Einzelfallbezug von Beratungen und ihre häufig vergleichsweise kurze Interventionsdauer erschweren es, Wirksamkeit in einen messbaren Rahmen zu

bringen. Es muss ebenfalls abgewogen werden, mit welchem Aufwand solche Untersuchungen durchgeführt werden. Für ein finanztechnisch „kleines“ Förderinstrument erscheinen umfassende Evaluierungen oder Wirkungsanalysen (beispielsweise Untersuchungen mit Kontrollgruppendesign) nicht angemessen. Nicht nur die Beratung selbst, sondern auch die Wirkungsmessung muss sich daher mit Effizienzfragen befassen (vgl. KÄPPLINGER 2014, S. 233).

Welche Möglichkeiten gibt es nun aber für Fördermittelgeber, mit begrenztem Aufwand eine valide empirische Grundlage zu schaffen, über die Förderinstrumente hinsichtlich ihrer Ergebnisse und ihrer Wirksamkeit bewertet werden können und die darüber hinaus die Basis für weitergehende Analysen der Instrumente bilden kann? Dieser Herausforderung wurde im Land Berlin mit der Entwicklung und Anwendung eines digitalen Monitoring- und Berichtssystems begegnet. Über das System werden Beratungsfalldaten einzelner öffentlich geförderter Beratungs- und Coachinginstrumente sowie im Rahmen einer Nachbefragung erhobene Angaben der Teilnehmenden erfasst, ausgewertet und in verschiedene Berichtsformate überführt. So werden insgesamt jedes Jahr über 100.000 Datensätze in Form von dokumentierten Konsultationen (durchgeführte Beratungen und Coachings) sowie Nachbefragungen über das System generiert. Durch die umfangreichen, systemseitigen Möglichkeiten der Datensynchronisierung für die statistische Auswertung werden für die einzelnen Förderinstrumente allgemeine Kennzahlen erzeugt und detaillierte Analysen durchgeführt. Die Ergebnisse der statistischen Auswertungen dienen sowohl den Beratungseinrichtungen zur Beobachtung, Steuerung und Bewertung auf der Mikroebene als auch einem periodischen Monitoring für ein begleitendes Fachcontrolling zum Instrumenteneinsatz durch den Fördermittelgeber, die Berliner Senatsverwaltung für Integration, Arbeit und Soziales, auf der Makroebene.

In diesem Beitrag aus der Praxis wird aufgezeigt, wie über ein solches System ein praktikables und systematisches Ergebnis- und Wirkungsmonitoring für öffentlich geförderte Beratungs- und Coachinginstrumente gewährleistet wird und wie eine Anschlussfähigkeit zur Evaluierung und Wirkungsforschung in der Berufsbildung hergestellt werden kann. Hierzu wird zunächst der aktuelle Forschungsstand zur Wirkungsmessung bildungs- und berufsbezogener Beratungsangebote skizziert. Anschließend werden die Funktionsweise und die Einsatzfelder des Berliner Monitoring- und Berichtssystems vorgestellt, um darauf aufbauend exemplarisch das Verfahren und die Resultate des Ergebnis- und Wirkungsmonitorings des Förderinstruments „Berliner Beratung zu Bildung und Beruf“ zu erläutern. Abschließend wird resümiert, wie das vorgestellte Verfahren bzw. die darüber erzeugten Daten für die Evaluierung und die Wirkungsforschung öffentlich geförderter Beratungs- und Coachingangebote im Bereich der beruflichen Bildung und der Arbeitsförderung besser genutzt werden können.

2 Wirkungsmessung von Beratung in wissenschaftlichen Studien

Auch wenn zunächst unter Wissenschaftlerinnen und Wissenschaftlern über den Nutzen Einigkeit besteht, Wirkung und Effektivität von Beratung messbar zu machen, muss gleichzeitig auf den geringen Umfang der Forschung in diesem Bereich hingewiesen werden (vgl. BIMROSE/BARNES 2009, S. 80).¹ Dies liegt sicher auch an den damit verbundenen Problemen. Die Schwierigkeiten, Wirkmechanismen zu identifizieren und die Wirkung von Beratung in einen messbaren Rahmen zu bringen, führt zu einer Uneinheitlichkeit der Qualitätssicherungssysteme und einer Uneinigkeit über zu verwendende Indikatoren (vgl. ebd., S. 82).

Einige Studien beschränken sich auf die Untersuchung „harter“ Wirkungsindikatoren, d. h. die Ermittlung von Erfolg anhand der konkret messbaren Beteiligung an Bildung, Arbeitsmarktintegrationen, Weiterbildung o. Ä. Das Europäische Netzwerk für eine Politik lebensbegleitender Beratung (*European Lifelong Guidance Policy Network*) hat im Rahmen seiner Bemühungen um die Erstellung eines Qualitätsrahmens bereits einige solcher Indikatoren vorgeschlagen. Darunter fallen z. B. Übergangsquoten in die nächsten Bildungs-, Ausbildungs- oder Beschäftigungsabschnitte, Verbleibsquoten und die Senkung von Abbruchquoten, das Ausmaß und die Dauer der Arbeitslosigkeit sowie eine erfolgreiche Abstimmung von Angebot und Nachfrage auf dem Arbeitsmarkt (vgl. SCHÖBER 2013, S. 32). Die Aussagekraft dieser Indikatoren allein ist jedoch begrenzt und wird auch als zunehmend inadäquat und problematisch bezeichnet, da sie nicht ausreichend die Bedürfnisse der Beratenden einbeziehen und somit nur zu einem Teil Beratungserfolg abbilden können (vgl. BIMROSE/BARNES 2009, S. 95; KÄPPLINGER 2010, S. 33). Es empfiehlt sich folglich, auch „weiche“ Indikatoren heranzuziehen, d. h. z. B. den Einfluss der Beratung auf die Motivation, auf die Handlungs- und Entscheidungsfähigkeit oder auch auf die Zufriedenheit der Ratsuchenden mit dem Resultat der Beratung.

Immer miteinbezogen werden muss zudem der Beratungsanlass sowie die persönliche Situation der Ratsuchenden. Damit eine Beratung als gelungen angesehen bzw. ihr Nutzen positiv eingeschätzt wird, kann auf den Zusammenhang zwischen den anfangs gesetzten Zielen und den erzielten Ergebnissen verwiesen werden (vgl. BEICHT/KREKEL/WALDEN 2004, S. 7). Die Teilnehmenden der Beratung unterscheiden sich in ihren persönlichen Merkmalen wie Geschlecht, Alter, Ethnizität sowie in ihren persönlichen Kontexten wie häusliche Situation, Familienstand und Wohnort (vgl. BIMROSE/BARNES 2009, S. 90). Solche verschiedenen Voraussetzungen führen wiederum zu unterschiedlichen Erwartungen und auch Abstufungen des Beratungserfolges. Zudem können persönliche Umstände wie Krankheit, strukturelle Arbeitslosigkeit oder Betreuungsverpflichtungen zu Beeinträchtigungen führen, die außerhalb der Kontrolle der Ratsuchenden liegen (vgl. ebd., S. 94). Bei der Bewertung des

1 Einen umfassenden Blick auf den Forschungsstand bietet das Handbuch zur Evidenzbasis lebensbegleitender Beratung des Europäischen Netzwerks für eine Politik lebensbegleitender Beratung (*European Lifelong Guidance Policy Network*). HOLLEY (2015) beschreibt darin die Wirkungen von Beratungen sowie die Methoden zu deren Überprüfbarkeit.

Beratungserfolges müssen solche Umstände in die Analyse miteinbezogen werden, damit es nicht zu Fehlschlüssen kommt.

In der Forschungsliteratur benannte Indikatoren beziehen sich auch auf die abschließende Bewertung der Beratung seitens der Ratsuchenden. Für den Erfolg von Beratung kann dabei auch die persönlich eingeschätzte, allgemeine Zufriedenheit von Bedeutung sein. Eine Unterscheidung der Zufriedenheit hinsichtlich verschiedener Bereiche (Beratung allgemein, Kompetenz des Beratenden, Beratungsergebnis, vgl. STROBEL 2010, S. 274) erscheint jedoch sinnvoll, um auch zielbezogen intervenieren zu können. Auch sollte die Befragung nicht nur unmittelbar nach dem Gespräch erfolgen, sondern nach einem längeren Zeitraum wiederholt werden, da bestimmte Umsetzungsschritte Zeit benötigen und der Erfolg der Umsetzung wiederum einen Einfluss auf die allgemeine Bewertung der Beratung hat (vgl. ebd., S. 276).

Wichtige Indikatoren, die auf die Entscheidungs- und Handlungsfähigkeit der Ratsuchenden zielen, sind eine größere Klarheit über Möglichkeiten und nächste Schritte, aber auch die Fähigkeit, die in der Beratung erhaltenen Informationen für das eigene Weiterkommen bewerten zu können (vgl. BIMROSE/BARNES/HUGHES 2008, S. 5). Manche Studien, vor allem aus dem englischsprachigen Raum, weisen zudem auf die Bedeutung der persönlichen Weiterentwicklung in Form von Selbstbewusstsein und einer erhöhten Sensibilisierung für persönliche Belange hin, welche als positives Ergebnis aus der Beratung entstehen können (vgl. BIMROSE/BARNES 2009, S. 93; KÄPPLINGER 2010, S. 34). Insgesamt werden durch diese Indikatoren die Kompetenzen und Fähigkeiten der Ratsuchenden verbessert, aber auch die Wahrnehmung für die eigenen Stärken geschärft, mithilfe derer die Ratsuchenden ihre persönlichen Bildungs- und Berufsziele erreichen können. Doch auch bei diesen verfeinerten methodischen Ansätzen bleibt die Frage einer kausalen Zuordnung der gemessenen Veränderungen zur Beratungsintervention eine wissenschaftliche und politische Herausforderung (vgl. SCHÖBER/LANGNER 2017, S. 19).

3 Zentrale Funktionsweisen und Einsatzbereiche des Berliner Monitoring- und Berichtssystems

Seit 2007 betreibt Arbeit und Leben (DGB/VHS) Berlin-Brandenburg ein onlinebasiertes Monitoring- und Berichtssystem. Der Entwicklung des Systems liegt eine Beauftragung der Berliner Senatsverwaltung für Integration, Arbeit und Soziales (SenIAS) zugrunde, ein integriertes Monitoringverfahren für verschiedene Landesinstrumente zur Beratung und zum Coaching im Bereich der Beruflichen Weiterbildung und der Arbeitsförderung zu entwickeln. Über das Monitoring sollen Kennzahlen ermittelt werden, welche die Basis für ein weiterführendes Fachcontrolling durch die SenIAS bilden. Das System wurde im Laufe der Jahre hinsichtlich seines Funktionsumfangs und seiner Einsatzfelder kontinuierlich weiter-

entwickelt.² Es wird mittlerweile als Software in Form einer onlinebasierten Arbeitsplattform angeboten und ermöglicht die Durchführung eines systematischen und kontinuierlichen Ergebnis- und Wirkungsmonitorings.

3.1 Funktionsweise

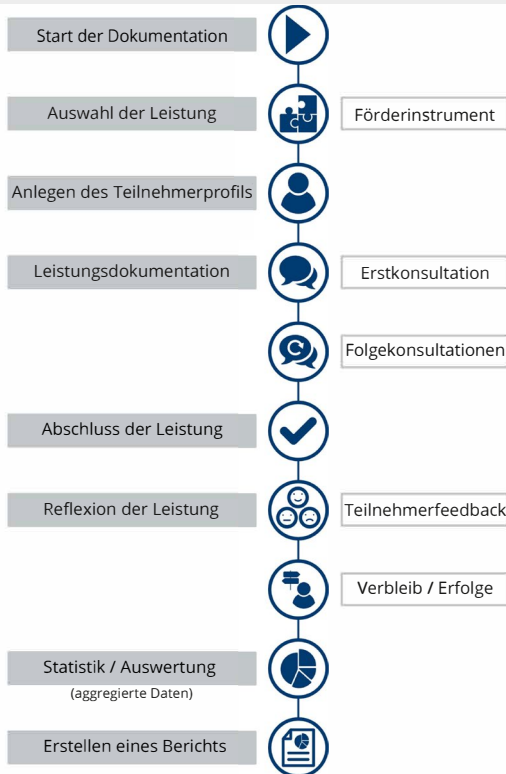
Das Monitoring- und Berichtssystem ermöglicht eine systematische und strukturierte Erfassung von Daten der Teilnehmenden und ihres Verbleibs, der erbrachten Beratungsleistungen sowie eine Onlinebefragung der Teilnehmenden. Die Daten werden über verschiedene Dokumentationsbögen erfasst. Die Beratenden nehmen Daten über die Beratung und den Beratungsprozess auf. Abschließend wird die Beratung von den Teilnehmenden beurteilt. Über das System lassen sich verschiedene Formen von Dokumentationsbögen (z. B. Teilnehmerprofil, Erst- und Folgekonsultation, Nachbefragung) und Abfragetypen (*Multiple Choice*, verschachtelte Abfragen, *Dropdown*-Auswahl, Freitextfelder) abbilden. Es ist somit geeignet, sowohl quantitative als auch qualitative Auswertungen von Beratungsprozessen zu generieren. Durch den modularen Aufbau kann eine individuelle Ausrichtung an unterschiedlichen Beratungsformaten berücksichtigt werden. In Abbildung 1 sind die verschiedenen Funktionsbereiche des Systems entlang eines typischen Anwendungsprozesses dargestellt.³

Über die in das System integrierte Echtzeitstatistik können die Akteure die für sie relevanten Ergebnisse in aggregierter Form generieren. Neben einer Reihe vordefinierter Kennzahlen (z. B. Dauer und Anzahl durchgeführter Beratungen, erbrachte Leistungen, Angaben zu den erreichten Zielgruppen) können sich die Nutzenden (Beratende, Projektleiter/-innen, Verantwortliche bei Auftraggebern und mittelverwaltenden Stellen) individuelle Statistiken zusammenstellen, in denen sich Daten beliebig miteinander in Beziehung setzen lassen. Die Statistiken können für beliebige Zeiträume ausgewiesen werden. Beim Anlegen einer Statistik wird zunächst definiert, ob ein einzelnes Merkmal abgebildet oder mehrere miteinander kombiniert werden sollen. Die Statistikfunktion im System ist mit einer Berichtsfunktion verknüpft, über welche die Einrichtungen den verantwortlichen Instanzen, gegenüber denen eine Berichtspflicht besteht, automatisch die benötigten Informationen zusenden können.

2 Seit dem Relaunch des Systems 2015 trägt es den Namen „Casian“ (von engl. „case“). Für detaillierte Informationen zum System vgl. www.casian.de.

3 Die Funktionen werden z. T. unterschiedlich in den Förderinstrumenten genutzt. So kommt das Teilnehmerfeedback in Form der automatisierten Nachbefragung nur in dem Förderinstrument „Berliner Beratung zu Bildung und Beruf“ zum Tragen (siehe hierzu Abschnitt 4).

Abbildung 1: Die Funktionsbereiche des Monitoring- und Berichtssystems



Quelle: eigene Darstellung

3.2 Einsatzbereiche

Das System ermöglicht ein systematisches Ergebnis- und Wirkungsmonitoring verschiedener Beratungs- und Coachinginstrumente. Beratende der von der SenIAS finanzierten großen Förderinstrumente „Berliner Beratung zu Bildung und Beruf“ (früher Bildungsberatung) und „Berliner Jobcoaching in der öffentlich geförderten Beschäftigung“ sowie kleinerer Instrumente nutzen das System. Zur Zeit werden jährlich Beratungs- bzw. Coachingdaten von ca. 36.000 Teilnehmenden dokumentiert und ausgewertet.

Der Fördermittelgeber nutzt das Ergebnis- und Wirkungsmonitoring für die Umsetzung eines Fachcontrollings, über das Transparenz erzeugt und die Umsetzung der Instrumente gesteuert werden soll (vgl. BERLINER SENATSVERWALTUNG FÜR INTEGRATION, ARBEIT UND SOZIALES 2017, S. 30). Die Trägereinrichtungen der Förderinstrumente können über das Ergebnis- und Wirkungsmonitoring die Umsetzung der Beratung bzw. des Coachings in ihrer

Einrichtung steuern und die Ergebnisse unmittelbar in ihr Qualitätsmanagement einfließen lassen.

In Abbildung 2 sind typische mit dem Ergebnis- und Wirkungsmonitoring verbundene Erkenntnisinteressen zusammengefasst abgebildet. Je nach Zielsetzung des Förderinstruments und auch in Abhängigkeit von der Rolle des Akteurs (Fördermittelgeber oder Träger-einrichtung) kann das Erkenntnisinteresse variieren.

Abbildung 2: Beispielhafte Erkenntnisinteressen eines Ergebnis- und Wirkungsmonitorings in der Beratung

Ebene	Inhalte	Fragen
 Prozess	Gestaltung, Methoden, Bedingungen	Wurden die Beratungsstandards erfüllt?
 Leistungen	Beratungsangebote, Veranstaltungen etc.	Wurden die angestrebten Fallzahlen erreicht?
 Nutzung	Sozialgruppen, Ausgangslagen, Bedarfe	Wurden die Zielgruppen erreicht?
 Wirkung	Fähigkeiten, Aktivitäten, Erfolge	Wurde das Besprochene umgesetzt?

Quelle: eigene Darstellung

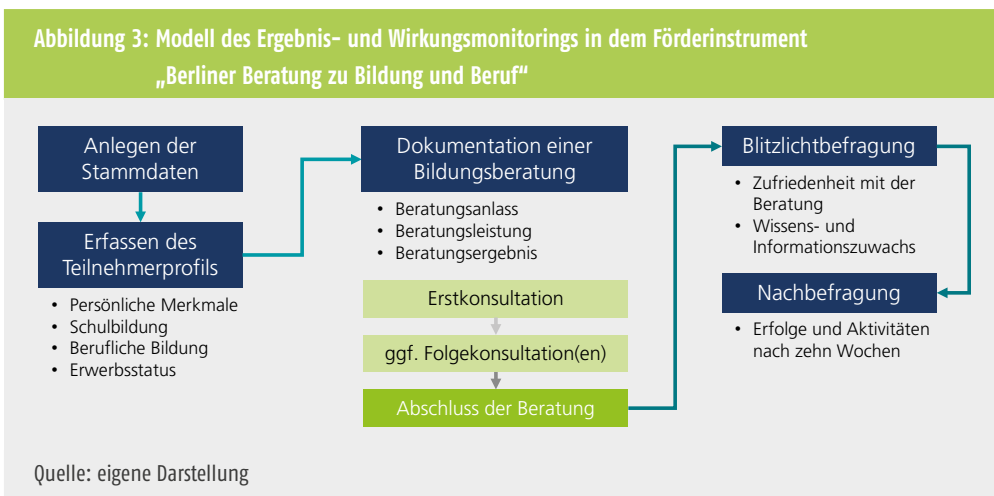
4 Das Ergebnis- und Wirkungsmonitoring am Beispiel der „Berliner Beratung zu Bildung und Beruf“ – Modell und ausgewählte Ergebnisse

4.1 Ansatz und Methode

Das für die „Berliner Beratung zu Bildung und Beruf“ gewählte Modell des Ergebnis- und Wirkungsmonitorings sieht die Erhebung von Indikatoren auf mehreren Ebenen vor. Die im Beratungsprozess durch die Berater/-innen erhobenen Daten werden nach der Beratung in einem zweistufigen Verfahren in Form von Nachbefragungen um die Perspektive der Beratenen ergänzt, sodass im Ergebnis ein Mehrebenenansatz für das Monitoring entsteht (vgl. KOMOSIN/KRUSE 2017, S. 78).

Die Dokumentation der Teilnehmenden und der Beratungsfalldaten durch die Beratenen ist dabei als Standardverfahren für die Vollerhebung konzipiert. Relevante Informationen müssen nicht an anderer Stelle dokumentiert werden. Dadurch werden prozessbegleitend Daten mit hoher Validität erhoben, und die Nachbefragung lässt sich effektiver

gestalten, da beispielsweise der Beschäftigungsstatus bereits erhoben wird und nicht mehr bei den Teilnehmenden abgefragt werden muss. Die Nachbefragung wiederum ist in zwei Schritte unterteilt. In der ersten Befragung („Blitzlicht“) werden die Zufriedenheit mit der Beratung sowie die Stärkung der Entscheidungs- und Handlungskompetenzen abgefragt, während es in der zweiten Befragung um die bildungs- und berufsbezogenen Aktivitäten und Erfolge geht. Die ausgefüllten Fragebögen werden über eine Schnittstelle in die Datenbank eingespeist, die Ergebnisse werden in aggregierter Form über die Onlinestatistik abgebildet. Die Wirkungsindikatoren können dabei mit den Merkmalen der Beratenen sowie den Beratungsdaten in Beziehung gesetzt werden. Abbildung 3 veranschaulicht das Modell des Ergebnis- und Wirkungsmonitorings:



4.2 Indikatorenbildung für das Wirkungsmonitoring

Die Basis des Wirkungsmonitorings bilden Ziele und sich daraus ableitende Indikatoren, welche geeignet scheinen, die Wirkungen und den Nutzen der Beratung für Bildung und Beruf abzubilden. Dem spezifischen Charakter dieses Beratungsangebots tragen zwei wesentliche Ziele Rechnung: Das erste Ziel ist gemäß OECD (2008), Menschen zu unterstützen, Bildungs- und Berufsentscheidungen auf einer gut vorbereiteten und informierten Basis eigenständig zu treffen und ihr Berufsleben selbst in die Hand zu nehmen. Das zweite Ziel leitet sich aus Arbeitsmarktprogrammen (hier das Programm „BerlinArbeit“ der SenIAS sowie das Operationelle Programm des Europäischen Sozialfonds (ESF) 2014–2020 für das Land Berlin) ab, die auf eine Erhöhung der beruflichen Chancen und der Bildungsbeteiligung abzielen.

Die Indikatoren geben Auskunft über die Zielerreichung. So werden im Rahmen der „Blitzlicht“-Befragung direkt nach der Beratung unmittelbare Wirkeffekte gemessen. Er-

hoben wird dies in zwei Fragekomplexen: Zum einen werden Fragen zur Zufriedenheit mit der Beratung als Grundbedingung für Beratungserfolg gestellt. Zum anderen werden über den Indikator „Wissens- und Informationszuwachs“ Hinweise generiert, inwieweit Entscheidungs- und Handlungskompetenzen gestärkt werden konnten. Operationalisiert wird dies durch Aussagen, die sich aus dem IOSM-Ansatz⁴ ableiten:

„Nach diesem Ansatz soll der Beratene anschließend informierter (in Bezug auf Möglichkeiten und Anwendung), orientierter (bezüglich eigener Potenziale und realistischer Ziele), strukturierter (in Bezug auf das Vorgehen und Erreichen der neuen Ziele), motivierter (bezüglich der eigenen Ziele im Verständnis von Eigensinn) sein“ (SCHLÖGL 2014, S. 140).

Über eine zweite Nachbefragung zehn Wochen nach der Beratung werden Indikatoren abgefragt, die sich auf die Umsetzung des Beratungsergebnisses beziehen, d. h. die mittelfristige Wirkung der Beratung. Sie bilden ab, inwieweit Beratung zur Beschäftigungsförderung und Erhöhung der Weiterbildungsbereitschaft beiträgt. Die Indikatoren „Teilnahme an Bildung“ und „Berufliche Entwicklung“ werden durch eine Reihe bildungs- und berufsbezogener Aktivitäten operationalisiert. Diese bilden ein Spektrum möglicher, für diesen Zeitraum realistisch scheinender Erfolge ab. Dem liegt die Annahme zugrunde, dass eine erfolgreiche Beratung nicht zwingend und häufig erst längerfristig in einen Statuswechsel mündet. So können z. B. bei Erwerbstätigen auch Verbesserungen bezüglich des Erhalts der Beschäftigungsfähigkeit erfasst werden.

Um die Aussagekraft der Indikatoren zu erhöhen, werden diese zu beiden Befragungszeitpunkten durch eine Selbsteinschätzung des Einflusses der Beratung auf die Erfolge ergänzt. Darüber hinaus können die Befragten offene Fragen beantworten.

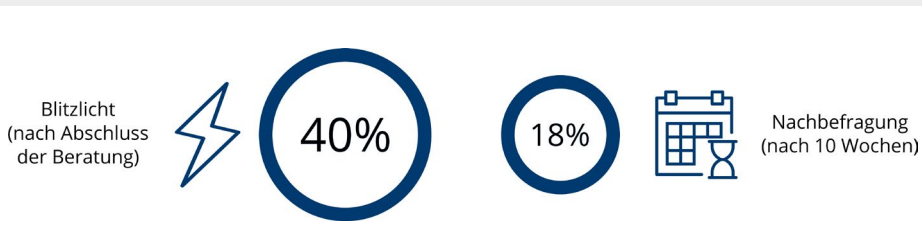
4.3 Ergebnisse des Wirkungsmonitorings aus 2017

Den folgenden Ergebnissen liegt die Auswertung der Fragebögen aus dem Jahr 2017 zugrunde, welche die Beratungsteilnehmenden unmittelbar nach Abschluss der Beratung („Blitzlicht“) sowie nach zehn Wochen (Nachbefragung) erhalten haben. Im Folgenden wird zuerst aufgezeigt, wie viele und welche Gruppen von Beratenen an der Befragung teilgenommen haben. Anschließend werden die Zufriedenheitswerte und die unmittelbaren Wirkeffekte sowie die mittelfristigen Erfolge und Aktivitäten nach zehn Wochen untersucht.

4 Das IOSM-Modell wurde von der k.o.s-GmbH als Bestandteil eines Qualitätskonzeptes für die Beratung zu Bildung und Beruf entwickelt und steht für „informierter, orientierter, strukturierter, motivierter“.

Rücklaufquote und Teilnehmendenstruktur

Abbildung 4: Rücklaufquote



Quelle: eigene Darstellung

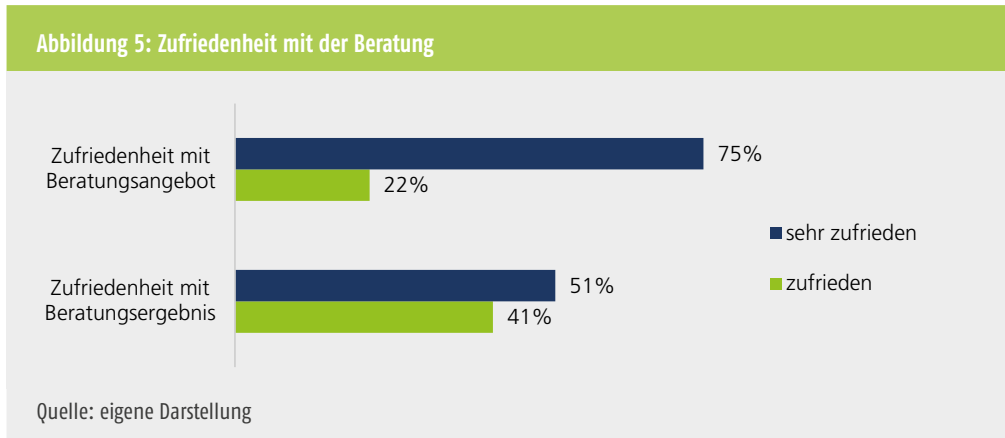
Im Jahr 2017 wurden von den Beratenen 882 Blitzlicht- und 466 Nachbefragungsbögen ausgefüllt. Die Rücklaufquote lag für die „Blitzlicht“-Befragung bei 40 Prozent und für die Nachbefragung bei 18 Prozent (siehe Abb. 4).⁵

Über die Personen, die an der Befragung teilgenommen haben, lassen sich folgende Aussagen treffen:

- ▶ Mit 78 Prozent (sowohl beim „Blitzlicht“ als auch bei der Nachbefragung nach zehn Wochen) haben prozentual mehr Frauen an der Befragung teilgenommen als insgesamt an der Bildungsberatung (67 %).
- ▶ Mit 27 Prozent beim „Blitzlicht“ bzw. 22 Prozent bei der Nachbefragung ist der Anteil von Personen mit Migrationshintergrund deutlich geringer als unter den Teilnehmenden der Bildungsberatung (47 %). Nicht ausreichende Deutschkenntnisse haben 13 Prozent („Blitzlicht“), während dies auf 29 Prozent der Migrantinnen und Migranten in der Bildungsberatung zutrifft.
- ▶ Mit 48 Prozent beim „Blitzlicht“ bzw. 52 Prozent bei der Nachbefragung ist der Anteil der zum Zeitpunkt der Beratung bereits Erwerbstätigen höher als unter den Teilnehmenden der Bildungsberatung insgesamt (37 %).

⁵ Die Fragebögen können von den Teilnehmenden erst ausgefüllt werden, nachdem ein Fall durch die Beratenen mit einer gültigen E-Mail-Adresse abgeschlossen wurde. Die Rücklaufquote errechnet sich aus dem Verhältnis von ausgefüllten Fragebögen zur Gesamtzahl der per E-Mail automatisch versandten Einladungen.

Zufriedenheit mit der Beratung



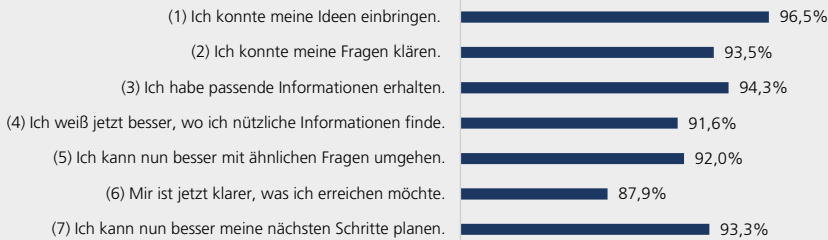
Es kann insgesamt eine hohe bis sehr hohe Zufriedenheit der Befragten festgestellt werden (siehe Abb. 5). Die Zufriedenheit mit dem Beratungsergebnis fällt regelmäßig etwas weniger hoch aus als die Zufriedenheit mit dem Beratungsangebot an sich (Beratungsstelle und Beratungspersonal). Mit unter fünf Prozent ist nur ein geringer Anteil der Teilnehmenden mit der Beratung weniger oder nicht zufrieden. Hinsichtlich der Gesamtzufriedenheit gibt es lediglich geringfügige Unterschiede zwischen verschiedenen soziodemografischen Gruppen, wobei Frauen, Personen ohne Migrationshintergrund und Nichterwerbstätige im Schnitt eine etwas höhere Zufriedenheit äußern.

Wissens- und Informationszuwachs

Der Gesamtwert für den Wissens- und Informationszuwachs ist ebenfalls sehr hoch. Er ermittelt sich aus dem Grad der Zustimmung für insgesamt sieben Aussagen.⁶ Bei einer Einzelbetrachtung der Aussagen lässt sich feststellen, dass Aussagen, die auf die Problemlösungsfähigkeit und die Umsetzung des Beratungsergebnisses zielen (4–7), tendenziell weniger stark zugestimmt wird als Aussagen, die eher auf die Beratungsqualität schließen lassen (1–3) (siehe Abb. 6). Männer sowie Personen mit Migrationshintergrund äußern etwas höhere Zustimmungswerte als der Durchschnitt der Befragten.

6 Die Zustimmung wird auf einer vierstufigen Skala abgefragt (ja – eher ja – eher nein – nein). Der Zustimmungswert in Prozent ergibt sich aus dem Anteil der Antworten mit „ja“ oder „eher ja“.

Abbildung 6: Wissens- und Informationszuwachs



Quelle: eigene Darstellung

Eine vertiefende Analyse der Daten aus der Beratung ergibt, dass die Zustimmungswerte bei Befragten, die zu Bildungsmöglichkeiten (geeignete Angebote/Anbieter, Finanzierungsoptionen u. Ä.) beraten wurden, etwas höher sind als im Durchschnitt. Im Schnitt zeigen sich niedrigere Zustimmungswerte, vor allem bei den Aussagen (5) und (6), bei Befragten, die zu bildungs- und berufsbezogenen Aktivitäten (Bilanzierung und Planung) oder zur Umsetzung von Bildungs- und Berufszielen beraten wurden.

Aktivitäten und Erfolge nach zehn Wochen

Abbildung 7: Aktivitäten und Erfolge nach zehn Wochen

Indikator
"Teilnahme an
Bildung"



22%

63%

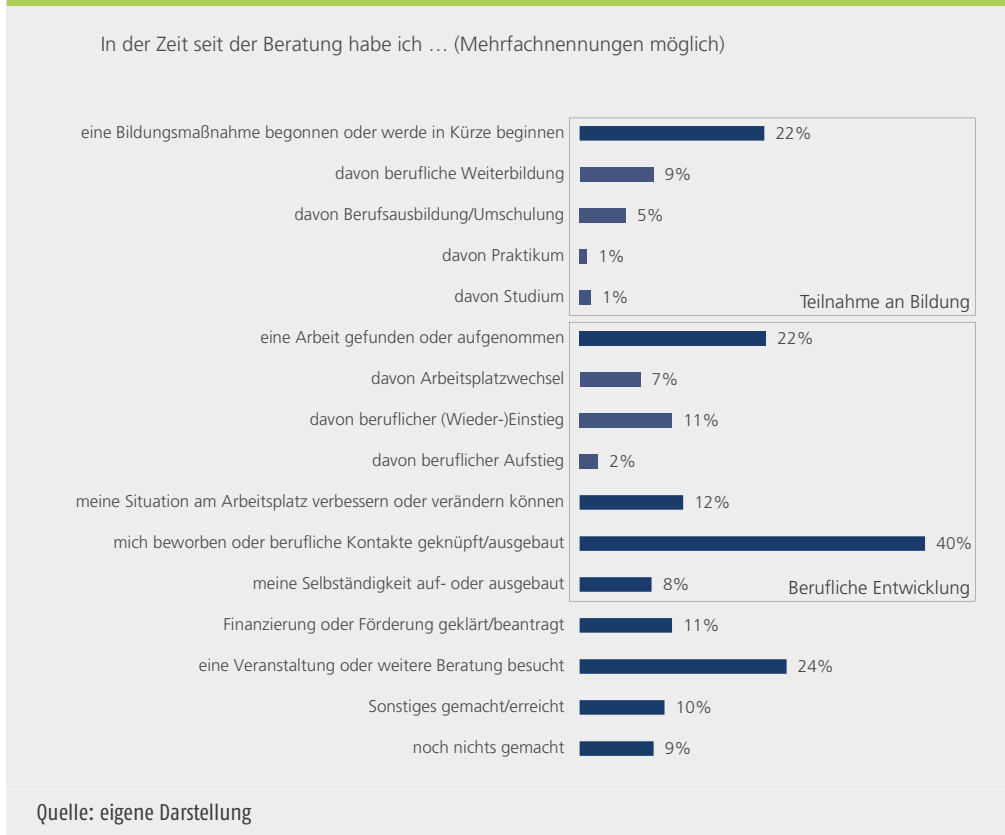


Indikator
"Berufliche
Entwicklung"

Quelle: eigene Darstellung

In der Nachbefragung nach zehn Wochen werden bildungs- und erwerbsbezogene Aktivitäten und Erfolge sichtbar gemacht. Insgesamt können 91 Prozent der Befragten mindestens eine Aktivität benennen, die darauf schließen lässt, dass innerhalb einer vergleichsweise kurzen Zeitspanne bereits konkrete Effekte zutage treten können. Lediglich neun Prozent der Befragten geben an, seit dem Abschluss der Beratung noch nichts unternommen bzw. erreicht zu haben (siehe Abb. 8).

Abbildung 8: Aktivitäten und Erfolge zehn Wochen nach der Beratung im Einzelnen



Aktivitäten und Erfolge, die eine berufliche Entwicklung aufzeigen, werden mit insgesamt 63 Prozent häufiger genannt als die Teilnahme an Bildung mit 22 Prozent oder weitere Aktivitäten. Eine Teilnahme an Bildung wird häufiger von Männern (24 %), Befragten mit Migrationshintergrund (26 %) sowie aktuell nicht erwerbstätigen Befragten (30 %) genannt. Die berufliche Entwicklung wird am häufigsten von zum Zeitpunkt der Beratung bereits Erwerbstätigen (68 % dieser Gruppe) genannt.

Eine vertiefende Analyse ergab, dass es eine hohe Übereinstimmung der nach zehn Wochen erzielten Ergebnisse mit den in der Beratung vereinbarten Folgeaktivitäten gibt. So gaben etwa 59 Prozent der Befragten, die in der Beratung die Klärung der Finanzierung einer Weiterbildung besprochen hatten, in der Befragung an, tatsächlich eine Bildungsmaßnahme begonnen zu haben. Von den Befragten, die in der Beratung Bewerbungsaktivitäten als nächste Schritte besprochen oder Bewerbungsgespräche vorbereitet hatten, gaben sogar 81 Prozent eine berufliche Entwicklung an.

5 Synergien erzeugen: Wie die Daten aus bestehenden Monitoringsystemen besser für die Evaluierung und Wirkungsforschung nutzbar gemacht werden können

Insgesamt stellt der hier vorgestellte Ansatz des Ergebnis- und Wirkungsmonitorings einen Versuch dar, mit einer langfristig umsetzbaren, praktikablen Lösung den kontinuierlichen Informationsbedürfnissen der beteiligten Akteure zu begegnen und erste Hinweise zur Wirksamkeit von Beratungs- und Coachingprozessen im Bereich der beruflichen Bildung und der Arbeitsförderung zu liefern. Durch die hohe Standardisierung in der Datenerfassung, die umfangreichen Möglichkeiten der Datensynchronisierung und durch die hohen Fallzahlen können sowohl verschiedene Ergebnis- und Wirkungskennzahlen erzeugt als auch detaillierte Analysen durchgeführt werden. Ein ähnliches – wenn auch in der technischen Umsetzung weniger ambitioniertes – System des Monitorings in der Berufsbildung wird über den ESF betrieben. Hier werden im Rahmen der Umsetzung von ESF-Maßnahmen umfassende Merkmale der Teilnehmenden sowie Ergebnis- und Zielindikatoren durch die Maßnahmenträger erhoben und zentral zusammengeführt (vgl. BAFZA 2020).

Sowohl die Berichtspflicht im Rahmen öffentlich geförderter Maßnahmen als auch die Notwendigkeit der evidenzbasierten Programm- bzw. Instrumentensteuerung setzen solche oder ähnliche Monitoringsysteme voraus. Um die dafür generierten Daten und Analysen über ihren eigentlichen Erzeugungskontext hinaus nutzbar zu machen, sollte geprüft werden, inwieweit sie in bestehende, auf Basis von amtlichen Statistiken erzeugte Systeme des Bildungs- oder Arbeitsmarktmonitorings einfließen könnten. Diese würden so durch einen „Beratungsmonitor“ ergänzt werden und ausgewählte Daten stünden einer breiten Fachöffentlichkeit zur Verfügung.⁷

Des Weiteren können die erzeugten Datenbestände für weitergehende Evaluierungen und die Wirkungsforschung genutzt werden. In Abhängigkeit von dem Erkenntnisinteresse einer Untersuchung kann der bestehende Datenbestand beispielsweise über Evaluierungsmethoden wie die Durchführung von Interviews bei Beratenden und Teilnehmenden gezielt um qualitative Daten ergänzt werden. Für Evaluierungen von Beratungsprozessen im Bereich der beruflichen Weiterbildung wären aus Sicht der Autorin und des Autors aktuell u. a. die folgenden Fragestellungen relevant:

7 Die Möglichkeiten der Einführung eines bundesweiten Beratungsmonitorings erörtern Schober und Käßlinger in ihrem Aufsatz „You get what you measure? – Überlegungen für ein Berichtssystem in der Bildungs- und Berufsberatung („Beratungsmonitor“)" 2017.

- ▶ Welche Hürden gibt es bei der Inanspruchnahme von Weiterbildungsmaßnahmen?
- ▶ Wie unterscheiden sich diese Hürden zwischen den verschiedenen Zielgruppen?
- ▶ In welcher Form spiegeln sich arbeitsmarkt- oder gesellschaftspolitische Trends (z. B. Digitalisierung, Grundbildung, prekäre Beschäftigung) in der Beratung und damit auch in der beruflichen Weiterbildung wider?
- ▶ Welche Beratungsbedarfe lassen sich hieraus ableiten?

Für tiefergehende Untersuchungen zur Wirksamkeit von Förderinstrumenten wiederum sollte geprüft werden, inwieweit – aufbauend auf einem bestehenden Ergebnis- und Wirkungsmonitoring – Untersuchungen mit Kontrollgruppendesign niedrigschwellig umgesetzt werden können. So ist es beispielsweise bei einem Coachinginstrument für Langzeitarbeitslose in einer öffentlich geförderten Beschäftigungsmaßnahme möglich, in der Untersuchung zwischen einer Gruppe mit Coaching und einer ohne Coaching zu unterscheiden.

Um die hier beschriebenen Synergien zwischen Monitoringsystemen in der beruflichen Bildung, der Arbeitsförderung und der Evaluierung und Wirkungsforschung herzustellen, bedarf es eines entsprechenden Impulses aus Politik und Wissenschaft. So könnten beispielsweise von der Politik entsprechende Forschungsaufträge formuliert werden, die explizit an vorhandene Monitoringsysteme anknüpfen. Für die Wissenschaft könnte es darüber hinaus eine Aufgabe sein, praxisorientiert aufzuzeigen, wie solche Monitoringsysteme noch besser für die Weiterentwicklung von Förderinstrumenten in der beruflichen Weiterbildung oder der Arbeitsförderung genutzt werden können.

Literatur

- BEICHT, Ursula; KREKEL, Elisabeth M.; WALDEN, Günther: Berufliche Weiterbildung – Welchen Nutzen haben die Teilnehmer? In: *Berufsbildung in Wissenschaft und Praxis* 33 (2004) 5, S. 5–9 – URL: <https://www.bwp-zeitschrift.de/de/bwp.php/de/bwp/show/936> (Zugriff: 28.07.2020)
- BERLINER SENATSVERWALTUNG FÜR INTEGRATION, ARBEIT UND SOZIALES (Hrsg.): Fachkonzept Beratung zu Bildung und Beruf. 2017 – URL: <https://www.berlin.de/sen/arbeit/weiterbildung/bildungsberatung/berliner-modell/> (Zugriff: 01.02.2019)
- BIMROSE, Jenny; BARNES, Sally-Anne: Measuring the effectiveness of career counselling in England: who defines what is effective? In: ARNOLD, Rolf; GIESEKE, Wiltrud; ZEUNER, Christine (Hrsg.): *Bildungsberatung im Dialog*. Band 2: 13 Wortmeldungen. III. Grundlagen der Berufs- und Erwachsenenbildung 61. Baltmannsweiler 2009, S. 79–98
- BIMROSE, Jenny; BARNES, Sally-Anne; HUGHES, Deirdre: *Adult Career Progression & Advancement: A Five Year Study of the Effectiveness of Guidance*. London/Coventry 2008

- BUNDESAMT FÜR FAMILIE UND ZIVILGESELLSCHAFTLICHE AUFGABEN (BAFzA) (Hrsg.): Monitoring und Evaluation 2014–2020. 2020 – URL: <https://www.esf-regiestelle.de/informationen/monitoring-und-evaluation.html> (Zugriff: 30.09.2020)
- HOOLEY, Tristram: Die Evidenzbasierung lebensbegleitender Beratung. Ein Wegweiser zu zentralen Forschungsergebnissen für eine effektive Politik und Praxis. In: EUROPEAN LIFELONG GUIDANCE POLICY NETWORK (Hrsg.): ELGPN Tools No. 3. Berlin 2015
- KÄPPLINGER, Bernd: Nutzen von Bildungsberatung. Konzeptionelle Eckpunkte vor dem Hintergrund britischer Forschungsergebnisse. In: DIE – Zeitschrift für Erwachsenenbildung 17 (2010) 2, S. 32–35
- KÄPPLINGER, Bernd; LICHT, Nina; RÄMER, Sabrina: „Ich sehe doch am Gesicht des Klienten, was wirkt und was nicht!“ In: MELTER, Ingeborg u. a. (Hrsg.): Zukunftsfeld Bildungs- und Berufsberatung III. Wirkung – Nutzen – Sinn. Bielefeld 2014, S. 213–238
- KOMOSIN, Birte; KRUSE, Henning: Der Mehrebenenansatz in der Ergebnis- und Wirkungsmessung für die Berliner Bildungsberatung. In: SCHÖBER, Karen; LANGNER, Judith (Hrsg.): Wirksamkeit der Beratung in Bildung, Beruf und Beschäftigung. Beiträge zur Wirkungsforschung und Evidenzbasierung. Bielefeld 2017, S. 77–90
- SCHLÖGL, Peter: Nutzen und Wirkung von Weiterbildungsberatung. In: SCHLÖGL, Peter; SCHRÖDER, Frank (Hrsg.): Weiterbildungsberatung. Qualität definieren, gestalten, reflektieren. In: Perspektive Praxis. Deutsches Institut für Erwachsenenbildung. Bielefeld 2014
- SCHÖBER, Karen: „Quality is a Journey, not a Destination“. Internationale Ansätze in Praxis, Politik und Wissenschaft zur Stärkung der Qualität und Professionalität in der Bildungs- und Berufsberatung. In: HAMMERER, Marika; KANELUTTI-CHILAS, Erika; MELTER, Ingeborg (Hrsg.): Zukunftsfeld Bildungs- und Berufsberatung II. Das Gemeinsame in der Differenz finden. Bielefeld 2013
- SCHÖBER, Karen; KÄPPLINGER, Bernd: „You get what you measure“? – Professionspolitische Überlegungen für ein Berichtssystem Bildungs- und Berufsberatung (Beratungsmonitor). In: SCHÖBER, Karen; LANGNER, Judith (Hrsg.): Wirksamkeit der Beratung in Bildung, Beruf und Beschäftigung. Beiträge zur Wirkungsforschung und Evidenzbasierung. Bielefeld 2017, S. 47–62
- STROBEL, Claudia: Bildungsberatung und Vernetzung auf kommunaler Ebene. Zentrale Ergebnisse einer Evaluationsstudie. In: GÖHLICH, Michael u. a. (Hrsg.): Organisation und Beratung. Wiesbaden 2010, S. 269–278

Andrea Mohorič, Nadja Konrad, Wolfgang Wittig

► **Evaluation des Förderprogramms „Berufsbildung ohne Grenzen“ zur betrieblichen Beratung bei der Auslandsmobilität von Auszubildenden und jungen Fachkräften**

Das Bundesministerium für Wirtschaft und Energie (BMWi) fördert auf der Grundlage der „Richtlinie zur betrieblichen Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften (Berufsbildung ohne Grenzen – BoG)“ insbesondere kleine und mittlere Unternehmen (KMU). BoG unterstützt dabei die Betriebe durch Mobilitätsberater/-innen, die bei Industrie- und Handelskammern sowie Handwerkskammern vor Ort ansässig sind. Sie beraten bei der Umsetzung von grenzüberschreitender Mobilität und der Antragstellung für den Bezug von Fördermitteln der Mobilitätsprogramme. Das im Beitrag dargestellte Evaluationsvorhaben bezieht sich auf die Vorgaben der Bundeshaushaltsordnung (BHO), die eine Erfolgskontrolle erforderlich macht. Auf Grundlage der Evaluationsergebnisse wurden Handlungsempfehlungen für die Fortführung des Förderprogramms formuliert.

1 Einordnung und Eckdaten zum Evaluationsauftrag

1.1 Einordnung

Die Strategie der Bundesregierung für die Internationalisierung von Bildung, Wissenschaft und Forschung macht deutlich, dass Deutschlands Bildungs-, Wissens- und Innovationssystem international ausgerichtet sein muss, um im weltweiten Wettbewerb bestehen zu können. Insbesondere im Bereich der beruflichen Aus- und Weiterbildung besteht daher ein großes Interesse an internationalen Kooperationen und der Entwicklung von internationalen beruflichen Handlungskompetenzen (vgl. BUSSE/FROMMBERGER 2016, S. 6). Aus diesem Grund spielt die Förderung der Auslandsmobilität der Auszubildenden und jungen Fachkräfte eine wichtige Rolle.

Die wachsende Bedeutung von Mobilität zeigt sich ebenfalls in der europäischen Bildungspolitik. Im europäischen Arbeitsprogramm „Allgemeine und berufliche Bildung 2020“ (erschienen 2020) wurde das Ziel definiert, die Mobilität der 18- bis 34-Jährigen mit einer beruflichen Erstqualifizierung bis zum Jahre 2020 auf sechs Prozent zu steigern. Auf nationaler Ebene wurde 2012 sogar das Ziel formuliert, dass mindestens zehn Prozent der Auszubildenden während der Ausbildung Auslandserfahrungen sammeln (vgl. BIBB 2017, S. 425). Der Datenreport zum Berufsbildungsbericht 2017 des Bundesinstituts für Berufsbildung zeigt einen stetigen Anstieg der Auslandsmobilität innerhalb des Programms Erasmus+ (2014–2020). Seit dem Jahr 2009 haben sich demnach die Förderzahlen der Auszubildenden und Berufsfachschüler/-innen mehr als verdoppelt, mit dem Ergebnis, das rund 21.000 Auszubildende und Berufsfachschüler/-innen durch Erasmus+ gefördert wurden. Unter der Annahme, dass die Zahl der Auslandsaufenthalte außerhalb von Erasmus+ seit 2010 stabil geblieben ist, so die Schätzung von Busse/Frommberger (2016), haben im Jahr 2016 mehr als 30.000 junge Menschen einen Auslandsaufenthalt im Rahmen der Erstausbildung absolviert, womit die Mobilitätsquote bei rund 5,3 Prozent liegt (vgl. BIBB 2018). Eine Steigerung auf zehn Prozent ist mutmaßlich nur durch zusätzlichen finanziellen Aufwand und strukturelle Maßnahmen zu erreichen (vgl. BUSSE/FROMMBERGER 2016, S. 87).

Die betrieblichen Fallstudien und Interviews, die von Busse/Frommberger durchgeführt wurden, zeigen, dass die Mobilitätsberatung ein wirkungsvolles Instrument zur Realisierung von Auslandsentsendungen von Auszubildenden sein kann. Weitergehend wird herausgestellt, dass es notwendig ist, den EU-Förderansatz durch entsprechende Beratungs- und Unterstützungsstrukturen zu begleiten, wenn in den nächsten Jahren der Anteil der Auszubildenden mit internationaler Erfahrung verdoppelt werden soll (vgl. BUSSE/FROMMBERGER 2016, S. 88).

1.2 Eckdaten

Seit dem 1. Januar 2015¹ fördert das BMWi die betriebliche Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften (BoG) insbesondere bei KMU.² Die administrative Durchführung der Förderung liegt beim Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA), das dem BMWi nachgeordnet ist.

BoG fördert die Beratung von Betrieben durch bei Industrie- und Handelskammern sowie Handwerkskammern vor Ort ansässige Mobilitätsberater/-innen. Die Betriebe werden bei der Umsetzung von grenzüberschreitender Mobilität und der Antragstellung von Fördermitteln der Mobilitätsprogramme unterstützt.

1 Vormalig wurde das Programm vom Bundesministerium für Arbeit und Soziales gefördert.

2 KMU sind Kleinunternehmen, kleine und mittlere Unternehmen, die laut EU-Regelung 2003/361 nicht mehr als 249 Beschäftigte haben und einen Jahresumsatz von höchstens 50 Millionen Euro erwirtschaften oder eine Bilanzsumme von maximal 43 Millionen Euro aufweisen (vgl. KOMMISSION DER EUROPÄISCHEN GEMEINSCHAFTEN 2003).

Nach einer Studie zur grenzüberschreitenden Mobilität in der Berufsbildung von 2011 sind deutsche Unternehmen zwar international aufgestellt, dennoch gehört die Entwicklung internationaler Berufskompetenzen der eigenen Fachkräfte nur in kleinem Umfang zu ihrer Ausbildungs- und Qualifizierungsstrategie. Denn nur etwa sieben Prozent aller Ausbildungsbetriebe entsenden regelmäßig oder gelegentlich Auszubildende ins Ausland (vgl. FRIEDRICH/KÖRBEL 2011, S. 62). In den Jahren 2014 bis 2017 haben laut der Mobilitätsstudie von 2017 nur 3,4 Auszubildende pro Betrieb einen Auslandsaufenthalt realisiert (vgl. NA BEIM BIBB 2017, S. 114).

Gerade KMU haben kaum Ressourcen, um Auslandsaufenthalte in der Erstausbildung zu organisieren und durchzuführen. Hier setzt das Programm BoG mit Beratung und Dienstleistung für Unternehmen an. Konkret werden derzeit 37 Mobilitätsberater/-innen bei 28 Trägern (fünf Industrie- und Handelskammern, 20 Handwerkskammern und drei freie Träger) und eine zentrale Koordinierungsstelle gefördert, die bei der Deutscher Industrie- und Handelskammertag (DIHK) Service GmbH und bei der Zentralstelle für Weiterbildung im Handwerk angesiedelt ist. Die Beratungen betreffen neben der Sensibilisierung von Unternehmen die Unterstützung der Betriebe bei der Organisation, Vorbereitung und Durchführung von Auslandspraktika und bieten Informationen zur Beantragung von Fördermitteln aus Mobilitätsprogrammen und die Begleitung der Antragstellung, z. B. bei Erasmus+. Sie ermöglichen kleinen und mittelständischen Betrieben Lernaufenthalte im Ausland. Zudem werden Betriebe auch bei der Aufnahme ausländischer Praktikanten und Praktikantinnen (*Incoming-Mobilität*) von den Mobilitätsberatern und -beraterinnen begleitet. Die zentrale Koordinierungsstelle bei der DIHK Service GmbH und der Zentralstelle für Weiterbildung im Handwerk unterstützt die Arbeit der Mobilitätsberater/-innen durch nationale und internationale Vernetzung, Öffentlichkeitsarbeit und Wissensmanagement sowie Monitoring und trägt so zu Kooperationen und Synergieeffekten bei der Umsetzung von BoG bei.

Das Evaluationsvorhaben des BMWi-Förderprogramms zur betrieblichen Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften bezieht sich auf die Vorgaben der BHO, die eine Erfolgskontrolle erforderlich machen.

Die Evaluation wurde von Januar bis Mai 2018 durchgeführt. Die Richtlinie des laufenden Programms zu BoG galt vom 1. Januar 2016 bis 31. Dezember 2019 (vgl. BAFA 2015b). Die Evaluation bezieht sich auf die Jahre 2016 und 2017.

Die programmegoischen Zielgrößen umfassen die Anzahl der durchgeführten Beratungen und Informationsveranstaltungen, die Anzahl der beratenen/teilnehmenden Betriebe und der durchgeführten Auslandspraktika.

2 Anforderungen und Bedingungen des Evaluationsauftrags

2.1 Untersuchungsziel und forschungsleitende Fragen

Die Erfolgskontrolle nach § 7 Abs. 2 BHO umfasst die drei Komponenten Zielerreichungs-, Wirkungs- und Wirtschaftlichkeitskontrolle. Die Evaluation bezieht sich auf insgesamt 14 Untersuchungsfragen (U1–U14), die vom Auftraggeber vorgegeben wurden. Die **Zielerreichungskontrolle** dient der Überprüfung, ob die mit dem Förderprogramm angestrebten Ziele erreicht wurden. Dieses Erkenntnisziel bezieht sich auf die in der Förderrichtlinie festgesetzten numerischen Zielgrößen. Zusätzlich wurde die Qualität der Beratung und Unterstützung betrachtet, die sich an der Zufriedenheit der Zielgruppen mit den Unterstützungs- und Beratungsleistungen bemisst. In diesem Kontext wurden auch die Qualität der Öffentlichkeitsarbeit, die Bekanntheit des Förderprogramms bei den Zielgruppen und die Nutzerfreundlichkeit der Webseite untersucht. Bei einer umfassenden Betrachtung des Untersuchungsgegenstands ist es erforderlich, neben den Zielgrößen und der Qualitätsdimension auch das Umfeld in den Blick zu nehmen, in dem diese Dienstleistungen erbracht werden. Darauf zielen die folgenden Untersuchungsfragen (U) der Evaluation ab:

- (U1) Wurden die angestrebten Zielgrößen des Programms BoG im Zeitraum von 2016 bis 2017 erreicht?
- (U2) Wie effektiv ist die Öffentlichkeitsarbeit? Ist das Förderprogramm bei den Zielgruppen ausreichend bekannt, wird die Webseite durch die Zielgruppen genutzt und wie wird ihre Nutzerfreundlichkeit beurteilt?
- (U3) Waren die KMU mit den Beratungs- und Unterstützungsleistungen der Mobilitätsberater/-innen zufrieden?
- (U4) Gibt es andere Initiativen und Projekte, die ähnliche Ziele verfolgen? Wenn ja:
 - ▶ Welches Alleinstellungsmerkmal hat BoG?
 - ▶ Wie nehmen die KMU die Attraktivität der BoG-Angebote im Vergleich zu anderen Initiativen und Projekten wahr?
 - ▶ Inwiefern ergänzen sich die verschiedenen Angebote? Inwiefern findet eine Kooperation statt?

Eine **Wirkungskontrolle** bezeichnet eine Analyse von Wirkungszusammenhängen bei der Umsetzung der Förderung unter dem Gesichtspunkt, ob und inwieweit die Förderung tatsächlich das Erreichen der Ziele bewirkt hat oder zumindest geeignet ist, diese zu bewirken. Dabei werden auf Basis der bisher betrachteten Erkenntnisse, Kontextfaktoren und vorhandene vergleichbare Angebote als Konkurrenz oder als sinnvolle Ergänzung untersucht. Zudem werden die Regionen, Branchen oder Berufsfelder der angesprochenen Unternehmen betrachtet. Zur Untersuchung der Wirkungskontrolle dienen die folgenden Fragestellungen:

- (U5) Inwieweit ist die Förderung von Beratungs- und Unterstützungsleistungen geeignet und ursächlich für das Erreichen der Ziele, d. h., ist eine Ausweitung und qualitative Verbesserung der Angebote zu verzeichnen, die nicht auch ohne die Förderung eingetreten wäre?
- (U6) Inwiefern hat die Förderung dazu geführt, dass Kammern sich zusätzlich in der Mobilitätsberatung engagieren?
- (U7) Inwiefern hat die Beratung dazu geführt, dass bei den beratenen KMU die Bereitschaft gestiegen ist, Auszubildende oder junge Fachkräfte ins Ausland zu entsenden? Führten die Erfahrungen mit realisierten Auslandsaufenthalten dazu, dass mehr KMU derartige Entscheidungen im Rahmen der Ausbildung vornehmen?
- (U8) Welche Faktoren fördern, welche hemmen die Umsetzung des Förderprogramms?
- (U9) Gibt es Unterschiede – im Hinblick auf die Wirksamkeit der Förderung – zwischen Regionen, Branchen oder Berufsgruppen?

Die **Wirtschaftlichkeitskontrolle** wird im Sinne der Vollzugs- sowie der Maßnahmenwirtschaftlichkeit durchgeführt. Zu prüfen ist, ob der beim Vollzug der Förderung, d. h. bei der Administration der Maßnahme, beim Erbringen der Beratungs- und Unterstützungsleistung durch die geförderten Einrichtungen und bei der zentralen Koordinierung und Steuerung des Angebots angefallene Ressourcenverbrauch erforderlich war. Zudem ist zu ermitteln, wie sich die realisierten Fördereffekte der Maßnahmen im Vergleich zu den angefallenen Kosten darstellen und ob diese Wirkungen auch mit niedrigeren Kosten hätten erzielt werden können. Das Verhältnis der erzielten Wirkungen zu den eingesetzten Ressourcen wird wesentlich durch die Ausgestaltung des Umsetzungsprozesses bestimmt. Für diese Analyse sind die folgenden Fragen relevant:

- (U10) Welche Personal- und Sachkosten sind bei der Erbringung der Beratungs- und Unterstützungsleistungen sowie der Koordination des BoG-Angebotes angefallen, und waren sie in dieser Höhe erforderlich?
- (U11) Gibt es bei der Durchführung und Administration der Förderung Verfahrensschritte, die einen Missbrauch begünstigen oder anderweitig zu ungerechtfertigten Kosten führen?
- (U12) Welche positiven Effekte aufseiten der KMU und der Auszubildenden bzw. Fachkräfte stehen den Kosten der Maßnahme gegenüber, und hätten diese Effekte auch mit geringeren Kosten erreicht werden können?

Auf der Grundlage der Ergebnisse der vorausgegangenen Struktur-, Prozess- und Ergebnisevaluation für den Zeitraum von 2016 bis 2017 werden **Handlungsempfehlungen** zu der Frage formuliert, ob und wie das Förderprogramm fortgeführt werden soll. Es werden mögliche Schwachstellen und Verbesserungspotenziale in der Durchführung der Prozesse identi-

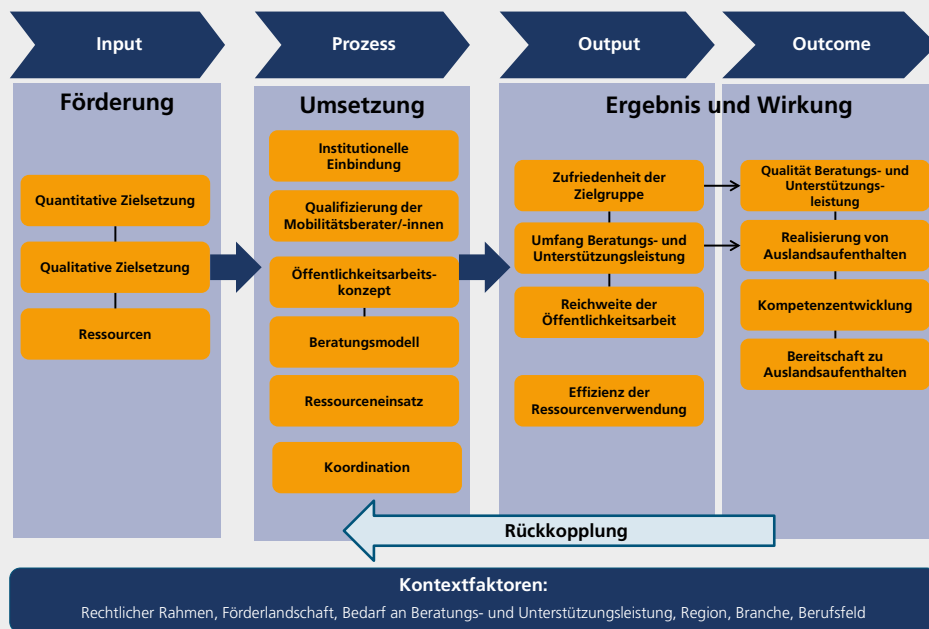
fiziert. Es geht um die Optimierung der Wirksamkeit der Förderung und die Gestaltung der organisatorischen Umsetzung, sodass Reibungsverluste möglichst vermieden werden.

- (U13) Was könnte verbessert werden, um die Effektivität des Förderprogramms zu erhöhen?
 (U14) Welche Änderungen im Verfahren könnten ggf. unternommen werden, um die Gefahr des Missbrauchs von Fördermitteln zu verringern?

2.2 Operationalisierung und Wirkungsmodell

Zur Operationalisierung der beschriebenen Fragestellungen wurde ein Wirkungsmodell formuliert. Das Modell unterscheidet zwischen Input, Prozess, Output und Outcome und nimmt des Weiteren auch die Kontextfaktoren in den Blick. Zweck dieses Modells ist es, die identifizierten Untersuchungsmerkmale zu ergänzen und zu systematisieren, um auf dieser Basis zu ersten Vermutungen über die bestehenden Wirkungszusammenhänge zu gelangen. Nachfolgend ist das Wirkungsmodell abgebildet und erläutert.

Abbildung 1: Übersicht über das Wirkungsmodell



Quelle: eigene Darstellung

Inputfaktoren sind die mit der Förderrichtlinie zur Verfügung gestellten (finanziellen) Ressourcen sowie die quantitative und qualitative Zielsetzung des Programms. Diese Inputfaktoren bestimmen den Handlungsspielraum der Mobilitätsberater/-innen und wirken sich damit auf die Zielerreichung aus.

Die Kategorie „**Prozess**“ beschreibt die organisatorischen, personellen und verfahrenstechnischen Umstände auf der Ebene der Mobilitätsberatung und auf der Ebene der zentralen Koordination. Die organisatorische und institutionelle Einbindung der Mobilitätsberater/-innen und ihre Qualifikation und Beratungsleistungen haben Einfluss auf die Effektivität der Leistungserbringung. Aber auch die Öffentlichkeitsarbeit, die Vernetzung der Mobilitätsberater/-innen mit weiteren nationalen und internationalen Akteuren, das Wissensmanagement und das Monitoring als Bestandteile der Tätigkeiten der Koordinierungsstelle beeinflussen die Effektivität der Leistungserbringung.

Auf der Ebene des **Outputs** zeigen sich die Auswirkungen der durchgeführten Beratungs- und Unterstützungsleistungen. Als Grundlage für die Bewertung können die Kennziffern des Monitorings, die Anzahl der durchgeführten Veranstaltungen, die Zahl der Entsendungen mit Betriebsbeteiligung und die Menge der Beratungen von Auszubildenden, jungen Fachkräften und Betrieben herangezogen werden. „Weiche“ Indikatoren wie die Effektivität der Öffentlichkeitsarbeit werden ebenso zur Dimension des Outputs gezählt.

Im **Outcome** spiegeln sich die Resultate der Förderung wider, d. h. die durch die Leistungen erzielten Wirkungen. Darunter wird die qualitative Dimension verstanden, z. B. die Zufriedenheit mit den erbrachten Beratungsleistungen. Der Erfolg kann sowohl anhand der Zahl der realisierten Auslandsaufenthalte als auch in einem Anstieg der Bereitschaft von KMU, ihre Auszubildenden und jungen Fachkräfte in ein Auslandspraktikum zu senden, bemessen werden.

Diese Faktoren in den einzelnen Dimensionen stehen in Wechselwirkung mit den **Kontextfaktoren**. Dabei handelt es sich um Einflussfaktoren außerhalb des Förderprogramms. Dazu gehören u. a. die Förderlandschaft, d. h. der Bestand an Projekten und Initiativen, die ein vergleichbares Beratungs- und Unterstützungsangebot zur Verfügung stellen, die Bedarfslage aufseiten der KMU und ihrer Beschäftigten, die regionale Wirtschaftsstruktur sowie der Stellenwert der internationalen Mobilität in den verschiedenen Branchen und Berufsfeldern.

3 Methodisches Vorgehen

3.1 Forschungsdesign

Methodisch wurde eine Kombination von quantitativen und qualitativen Forschungsverfahren angewendet. Qualitative Methoden dienen in diesem Fall der vertieften Untersuchung der verschiedenen Aspekte der Mobilitätsberatung. Quantitative Forschungsmethoden liefern belastbare Ergebnisse und ermöglichen Schlussfolgerungen im Hinblick auf die Wirksamkeit und Wirtschaftlichkeit der Förderung.

Die Untersuchung der Wirksamkeit des Programms basiert auf einem Kontrollgruppen-design, bei dem folgende Vergleichsgruppen außerhalb des Förderprogramms betrachtet wurden:

Vergleichsgruppe „Andere Mobilitätsberater/-innen“: Hinsichtlich der Beratungsarbeit und der Beratungserfolge werden sowohl im Rahmen der Dokumentenanalyse als auch der qualitativen Interviews folgende Vergleichsgruppen betrachtet oder befragt:

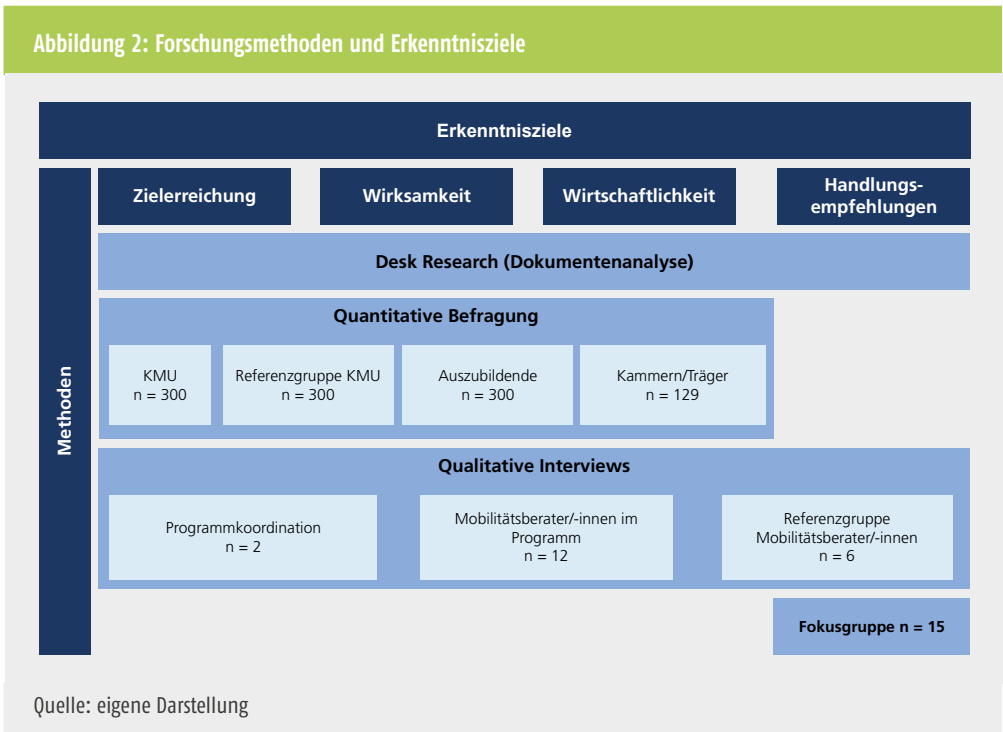
- ▶ Mobilitätsberater/-innen ohne BoG-Förderung, die bei den Kammern und anderen Trägern angesiedelt sind;
- ▶ Mobilitätsberater/-innen der Informations- und Beratungsstelle für Auslandsaufenthalte in der beruflichen Bildung (IBS bei der Nationalen Agentur Bildung für Europa beim BIBB);
- ▶ Mobilitätsberater/-innen der Nationalen Agentur Bildung für Europa beim Bundesinstitut für Berufsbildung (NA beim BIBB);
- ▶ Berufsschullehrer/-innen, die Mobilitätsberatungen an berufsbildenden Schulen durchführen; dabei sollen ausgewählte Bundesländer mit hoher Mobilitätsquote betrachtet werden.

Vergleichsgruppe „Andere Betriebe“: Hinsichtlich der Referenzgruppen zu beratenen Betrieben und deren Erfahrungen wird sowohl im Rahmen der Dokumentenanalyse als auch im Rahmen der quantitativen Befragungen folgende Vergleichsgruppe analysiert und befragt:

- ▶ Betriebe, die keine Mobilitätsberatung im Rahmen des BoG-Förderprogramms erhalten oder keine bzw. eine andere Mobilitätsberatung in Anspruch genommen haben.

Die Zuordnung der verschiedenen Methoden zu den Erkenntniszielen ist in Abbildung 2 dargestellt.

Abbildung 2: Forschungsmethoden und Erkenntnisziele



3.2 Methoden

Desk Research/Dokumentenanalysen

Zu Beginn der Forschungsarbeiten erfolgte die Auswertung von Forschungsliteratur und weiteren schriftlichen Quellen zum Thema. Im Fokus standen die vorliegenden Daten zur Anzahl und Verteilung der durchgeführten Beratungsaktivitäten und Informationsveranstaltungen. Ziel war es, mit Blick auf die Zielerreichungs- und Wirkungskontrolle Grundlagen für die Erschließung und Strukturierung des Feldes zu ermöglichen und Hypothesen zu den Wirkungszusammenhängen auszuformulieren. Darüber hinaus wurden weitere Anhaltspunkte für die Beurteilung der Wirtschaftlichkeit – vor allem im Hinblick auf die Auswahl geeigneter Nutzenindikatoren zur Prüfung der Maßnahmenwirtschaftlichkeit – sowie erste Hinweise auf mögliche Entwicklungspotenziale und Handlungsempfehlungen gewonnen. Bei den Analysen wurden z. B. Monitoringberichte, Materialien der Öffentlichkeitsarbeit der Koordinierungsstelle und einzelne Mobilitätsberater/-innen einbezogen.

Für die Wirtschaftlichkeitskontrolle, die im vorliegenden Beitrag aus Platzgründen nicht weiter behandelt wird, wurden Unterlagen zu den geförderten Mobilitätsberatungsprojekten (im Folgenden als „Projekte“ bezeichnet), d. h., Anträge und Sachberichte, beim BAFA analysiert. Dafür wurden eine Analysematrix angelegt und Leitfragen für die Untersuchung des Programms entwickelt. Zudem wurden vom BAFA zur Verfügung gestellte Aufstellungen

der bewilligten und abgerufenen Fördermittel sowie die beim BAFA selbst im Zusammenhang mit der Programmdurchführung angefallenen Kosten ausgewertet. Eine weitere Datengrundlage bildete eine Stichprobe von fünf geförderten Projekten, die Zahlen zu geplanten Personal-, Sach-, und Reisemitteln und Informationen zur Zielerreichung bei Beratungen, Veranstaltungen und Mobilitäten mit Betriebsbeteiligung zur Verfügung stellt. Zudem wurden Materialien der Öffentlichkeitsarbeit, die durch die Koordinierungsstelle erstellt werden, exemplarisch auf ihre Kosten und Verbreitung hin geprüft.

Standardisierte Befragung

Die Perspektiven der Zielgruppen des Förderprogramms (KMU und Auszubildende bzw. junge Fachkräfte) sowie der Kammern auf die Umsetzung der Mobilitätsberatung und ihre Auswirkungen wurden durch standardisierte Befragungen erhoben. Die Onlinebefragung wurde mit der Befragungssoftware LimeSurvey durchgeführt und über den Server des Forschungsinstituts Betriebliche Bildung – f-bb zur Verfügung gestellt.

Die Internetlinks zu den Onlinefragebögen für die teilnehmenden Betriebe und Auszubildenden wurden aus Datenschutzgründen zentral an die Mobilitätsberater/-innen im Förderprogramm versendet. Ein Unterstützungsschreiben vom BMWi wurde zur Motivation der Befragten beigefügt. Die Mobilitätsberater/-innen haben die Links zur Onlinebefragung an die von ihnen beratenen Betriebe und Auszubildenden weitergeleitet. Es gab Mobilitätsberater/-innen, welche die von ihnen beratenen Betriebe und Auszubildenden besonders zur Teilnahme motiviert haben, sodass diese besonders stark bei der Befragung vertreten sind. Ein Nachfassen, z. B. im Hinblick auf die Verteilung der Bundesländer, war nicht möglich. Es ließ sich auch nicht kontrollieren, wer von den Angesprochenen die Onlinebefragung tatsächlich weitergeleitet hat. Eine Gewichtung der Daten der teilnehmenden Betriebe, der Referenzbetriebe und Auszubildenden ist nicht möglich, da nicht bekannt ist, welche prozentuale Verteilung von ihnen nach Region und Branche in den beteiligten Kammerbezirken vorliegt. Deshalb kann die Untersuchungsfrage 9 „Gibt es Unterschiede – im Hinblick auf die Wirksamkeit der Förderung – zwischen Regionen, Branchen oder Berufsgruppen?“ nur unvollständig ausgewertet werden. Die vorliegenden Daten lassen keine verlässlichen Angaben zu.

Durch die Ansprache der Betriebe und Auszubildenden über die Mobilitätsberater/-innen war die Auswahl dem Zufall des Rücklaufs überlassen. Im Fall der teilnehmenden Betriebe wurde eine Stichprobe von n=238 beratenen Unternehmen befragt; als Kontrollgruppe dienten Unternehmen, die das Beratungsangebot des Programms BoG nicht in Anspruch genommen haben (n=189 Unternehmen).

Die Kontaktdaten der Referenzgruppe der Betriebe wurden nach Regionen und Branchen gezielt recherchiert und beschafft. Dabei wurde der Fokus auf Betriebe in Regionen und Branchen gelegt, die im Förderprogramm bisher nicht bzw. eher wenig vertreten sind, damit ermittelt werden kann, warum diese nicht teilnehmen bzw. welche Unterstützung diese Betriebe von der Mobilitätsberatung für eine Teilnahme brauchen. Die Ansprechpartner/-innen in den Kammern wurden manuell recherchiert und mithilfe eines Unterstützungsschrei-

bens der Dachverbände DIHK und Zentralverband des Deutschen Handwerks gemeinsam mit dem BMWi zur Teilnahme motiviert.

Im Kick-off-Gespräch mit dem Auftraggeber der Evaluation wurde das Evaluationskonzept angepasst und eine Vollerhebung mit 93 teilnehmenden Kammern realisiert. Der Grund dafür war der Wunsch, Faktoren zu erfassen, welche die Kammern von einer Teilnahme am Programm abhalten. Das zukünftige Ziel des BMWi ist es, weitere Kammern für die Teilnahme am Förderprogramm zu gewinnen.

Die angestrebten Stichprobenzahlen von 300 Unternehmen, 300 Referenzunternehmen und 131 Kammern konnten trotz zweimaligem Nachfassen nicht erreicht werden. Gründe dafür lagen vermutlich in der Befragungszeit während der Oster- und Pfingstferien. In Absprache mit dem Auftraggeber wurden die erreichten Anzahlen als ausreichend betrachtet.

Insgesamt haben 427 Unternehmen aus allen Bundesländern, ausgenommen Saarland, an der Befragung teilgenommen. Die befragten Betriebe sind vor allem aus der Branche des Baugewerbes (teilnehmende Betriebe 19,7 %, Referenzbetriebe 24,1 %), des verarbeitenden Gewerbes (teilnehmende Betriebe 19,7 %, Referenzbetriebe 11,8 %) und des Handels, der Instandhaltung und Reparatur von Kraftfahrzeugen (teilnehmende Betriebe 10,3 %, Referenzbetriebe 18,2 %).³

Innerhalb der Kammern wurde der Onlinefragebogen zu 50 Prozent von Abteilungs- bzw. Bereichsleiter/-innen ausgefüllt. 16,7 Prozent der Kammerbefragten, die den Onlinefragebogen ausgefüllt haben, sind aus der Abteilung der Ausbildung bzw. (Aus-)Bildungsberatung, 13,3 Prozent in der Abteilung der Bildung/Berufsbildung und je zehn Prozent in der Abteilung Aus- und Weiterbildung ihrer Kammer bzw. sind selbst Mobilitätsberater/-innen.⁴

Insgesamt haben 300 Auszubildende aus 13 Bundesländern an der Befragung teilgenommen, damit wurde die angestrebte Stichprobengröße realisiert. 52,8 Prozent der befragten Auszubildenden sind Frauen und 46,8 Prozent der befragten Auszubildenden sind Männer, 0,3 Prozent haben keine Angabe zu ihrem Geschlecht gemacht.⁵ Die meisten der befragten Auszubildenden arbeiten im Baugewerbe (17,3 %), im verarbeitenden Gewerbe (14,5 %) und in sonstigen Dienstleistungsberufen (11,0 %).⁶

83,2 Prozent der befragten Auszubildenden haben bereits ein Auslandspraktikum absolviert, 14,3 Prozent von ihnen haben ein Auslandspraktikum geplant und 2,5 Prozent haben ein Auslandspraktikum weder absolviert noch geplant.⁷

Qualitative Interviews

Die Perspektive der Akteure im Förderprogramm BoG wurde durch qualitative Interviews erfasst. Es handelte sich jeweils um leitfadengestützte Einzelinterviews, die einer qualitativen

3 Insgesamt haben 223 teilnehmende Betriebe und 170 Referenzbetriebe auf diese Frage geantwortet.

4 Insgesamt haben 90 Kammern auf diese Frage geantwortet.

5 Insgesamt haben 299 Auszubildende auf diese Frage geantwortet.

6 Insgesamt haben 283 Auszubildende auf diese Frage geantwortet.

7 Insgesamt haben 280 Auszubildende auf diese Frage geantwortet.

Inhaltsanalyse unterzogen wurden. Die Interviews, die mit Akteuren sowohl auf der Ebene einzelner geförderter Einrichtungen als auch auf der Ebene der zentralen Programmkoordination geführt wurden, dienten der Erhebung von Informationen zu drei Erkenntniszielen, d. h. der Zielerreichungs-, Wirtschaftlichkeits- und Wirkungskontrolle.

Die Zielpersonen der qualitativen Interviews waren die Vertreter/-innen der Koordinierungsstelle (n=2) und die Mobilitätsberater/-innen, die im Programm BoG gefördert werden (n=12) und von denen vier auf Industrie- und Handelskammern, sieben auf Handwerkskammern und eine auf eine kammerunabhängige Mobilitätsberatung entfallen. Des Weiteren wurden sechs Interviews mit Mobilitätsberatern und -beraterinnen außerhalb des Förderprogramms geführt, um das „Kontrollgruppendesign“ auch auf der Ebene der Akteure zu realisieren. Von diesen Interviewpersonen sind je eine bei der NA beim BIBB und bei der IBS und je zwei an Kammern außerhalb des Förderprogramms bzw. an Schulen tätig.

Die Interviews umfassten folgende Themen:

- ▶ Arbeitsschwerpunkte und Selbstverständnis der Beratung,
- ▶ Förderprogramm,
- ▶ Qualitätssicherung durch Weiterbildung/Einarbeitung und Monitoring/Kennzahlen,
- ▶ Öffentlichkeitsarbeit,
- ▶ Beratung von Betrieben/Betriebsakquise,
- ▶ Rolle und Unterstützung der Kammern,
- ▶ Kooperationen mit anderen Projekten/Initiativen,
- ▶ Austausch im Netzwerk.

Fokusgruppe

Zur Validierung der Ergebnisse und zur Unterstützung der Ableitung von Handlungsempfehlungen wurde am 8. Mai 2018 abschließend eine Fokusgruppe mit 15 Akteuren der verschiedenen Ebenen des Förderprogramms durchgeführt (Vertreter/-innen der Industrie- und Handelskammern, Handwerkskammern und deren Dachorganisationen, Vertreter/-innen der Organisationen, bei denen die zentrale Koordinierungsstelle angesiedelt ist, Vertreter/-innen von Alternativangeboten). Das Verfahren der Gruppendiskussion bietet sich zur Präzisierung der Handlungsempfehlungen an, da es sich bei den teilnehmenden Akteuren um relativ klar abgegrenzte Kollektive oder Milieus handelt, in denen sich bereits gemeinsame Auffassungen und Deutungsmuster herausgebildet haben, die durch den Diskussionsprozess nicht erst erzeugt, sondern lediglich aktualisiert werden (vgl. BOHNSACK 2009, S. 370–374).

Die Teilnehmenden der Fokusgruppe diskutierten vorab versendete Handlungsempfehlungen, die aus Zwischenergebnissen der Experteninterviews und der Onlinebefragungen abgeleitet wurden. Während der Diskussion wurden die Handlungsempfehlungen von den

Teilnehmenden ergänzt/präzisiert, bestätigt/verstärkt oder verworfen bzw. andere Handlungsempfehlungen wurden angebracht.

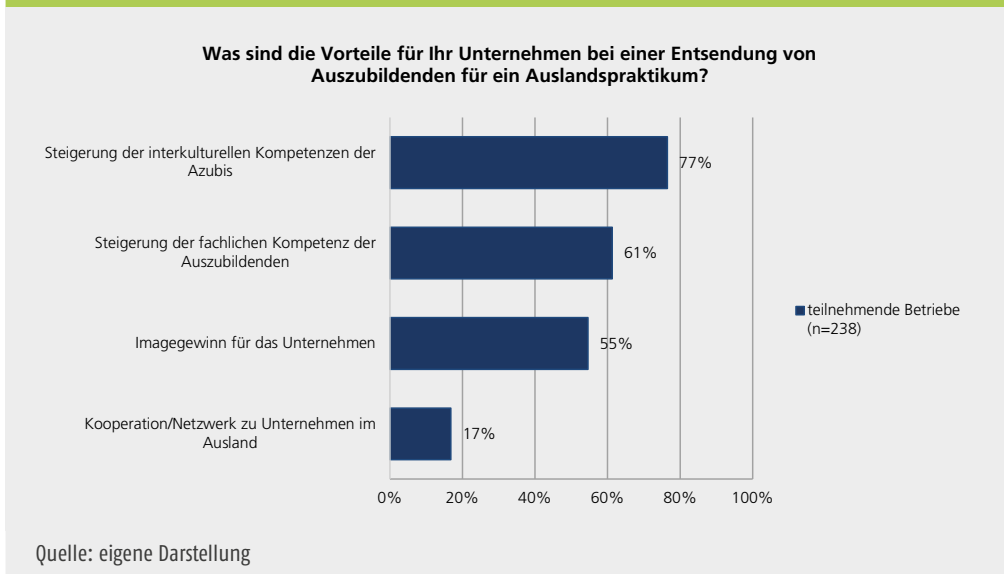
4 Ausgewählte Ergebnisse

4.1 Erfolge und Stärken der Mobilitätsberatung

Die Mobilitätsberater/-innen des Förderprogramms leisten sehr gute und wichtige Arbeit. Dies wird vor allem an den Ergebnissen aus der Onlinebefragung der teilnehmenden Betriebe deutlich. So hätten 78 Prozent der befragten Betriebe ein Auslandspraktikum nicht ohne eine Mobilitätsberatung durchgeführt⁸, und 85 Prozent gaben an, weiterhin Auszubildende zu entsenden.⁹

Der Nutzen der Auslandsmobilität für Betriebe wurde bei Betrieben erhoben, die in den letzten zwei Jahren Auszubildende entsandt haben. Die Stichprobe umfasste 238 teilnehmende Betriebe. Die Betriebe wurden nach den Vorteilen bei der Entsendung von Auszubildenden für ein Auslandspraktikum gefragt. 77 Prozent gaben die Steigerung der interkulturellen Kompetenz der Auszubildenden an, 61 Prozent die Steigerung der fachlichen Kompetenz der Auszubildenden und 55 Prozent den Imagegewinn für das Unternehmen. 17 Prozent sahen den Vorteil in der Kooperation zu Unternehmen im Ausland (siehe Abb. 3).

Abbildung 3: Vorteile für die Unternehmen bei einer Entsendung von Auszubildenden



8 Insgesamt haben 207 teilnehmende Betriebe auf diese Frage geantwortet.

9 Insgesamt haben 211 teilnehmende Betriebe auf diese Frage geantwortet.

Insgesamt wird deutlich, dass der Mehrwert eines Auslandspraktikums mittlerweile über die bloße Entwicklung des einzelnen Auszubildenden hinausgeht. So wird eine Außenwirkung des eigenen Unternehmens von einigen Betrieben bereits mitgedacht. Ausbildungsmarketing und die Kooperation mit Betrieben und Partnern im Ausland werden zunehmend ein Thema.

Bei der Befragung der teilnehmenden Betriebe zu den Auswirkungen der Mobilität auf die Auszubildenden wird vor allem von der persönlichen Reife und dem gestärkten Selbstvertrauen berichtet (78 %). Von einem Einblick der Jugendlichen in andere Lebens- und Arbeitswelten sowie vom Kennenlernen anderer Verfahren, Abläufe und Arbeitsweisen berichten je 66 Prozent der Betriebe. Auch der Zuwachs an Kompetenzen wird von vielen befragten Unternehmen betont: 62 Prozent berichten von gesteigener Sozialkompetenz, 61 Prozent von einer erweiterten Sprachkompetenz sowie 54 Prozent von einer erweiterten interkulturellen Kompetenz und 44 Prozent geben eine gestiegene Fachkompetenz an (vgl. Abb. 4).¹⁰

Dies deckt sich mit dem Befund der aktuellen Mobilitätsstudie der NA beim BIBB, wonach der Nutzen eines Auslandsaufenthaltes nach dem übereinstimmenden Urteil der befragten Auszubildenden, Betriebe und Schulen vornehmlich in der Entwicklung der persönlichen Kompetenzen liegt (vgl. NA BEIM BIBB 2018, S. 99–101). Hierbei liegt das Selbstbewusstsein mit einem Mittelwert von 1,9 auf einer Skala, bei der 1 die stärkste und 5 die schwächste Ausprägung eines Merkmals repräsentiert, an der Spitze, gefolgt von Verantwortungsbereitschaft, Motivation und selbstständigem Arbeiten. Ebenso wird auch in der Mobilitätsstudie ein noch verhältnismäßig starker positiver Effekt auf die Sprachkenntnisse verzeichnet (Mittelwert 1,8 aus Sicht der Betriebe und Schulen), während die Förderung der Fachkompetenz mit Mittelwerten um 2,5 schwächer beurteilt wird (vgl. NA BEIM BIBB 2018, S. 102–103).

Abbildung 4: Vorteile des Auslandsaufenthalts für die Auszubildenden



10 Insgesamt haben 238 teilnehmende Betriebe auf diese Frage geantwortet.

Die positiven Lerneffekte der Auslandsmobilität münden in die internationale berufliche Handlungskompetenz, die Frommberger in der Studie „Internationalisierung der Berufsausbildung als Herausforderung für Politik und Betriebe“ (2016) als wesentliche Kompetenz hervorhebt, um den Anforderungen durch die Internationalisierung der Wirtschafts-, Arbeits- und Lebenswelt zu begegnen: „Allgemein lässt sich berufliche Handlungskompetenz als ein Bündel an Kenntnissen, Fertigkeiten und Einstellungen beschreiben, das seinen Träger in die Lage versetzt, sich in einem internationalen beruflichen Umfeld zu bewegen“ (BUSSE/FROMMBERGER 2016, S. 28).

Zusammenfassend sehen die befragten Unternehmen bei der Entsendung von Auszubildenden für ein Praktikum ins Ausland primär den Nutzen der Persönlichkeitsentwicklung ihrer Auszubildenden. Die Erweiterung von deren fachlichen und sozialen Kompetenzen ist für die Betriebe ebenfalls ein wichtiges Ergebnis des Auslandspraktikums.

Die Motive für Auslandsmobilität liegen für Betriebe also darin, Auszubildende zu gewinnen, Fachkräfte zu binden und die internationale Geschäftstätigkeit auszubauen. In den Experteninterviews und der Fokusgruppe wurde mehrfach geäußert, dass Ausbilder/-innen vor allem dann überzeugt sind, wenn sie selbst einen Auslandsaufenthalt erlebt und begleitet haben. Zudem zeigt sich, dass ein Betrieb, der einen Auszubildenden einmal entsendet hat, auch weiterhin Mobilität unterstützt. Dies könnte genutzt werden, um erfahrene Unternehmen als Botschafter – aktive Akteure – im Förderprogramm BoG einzusetzen.

4.2 Überzeugung neuer Betriebe und Alleinstellungsmerkmal verstärken

Deutlich werden die Vorteile der Mobilitätsberatung für die Auslandsmobilität bei teilnehmenden Betrieben. Allerdings ist BoG nur bei 13 Prozent der Referenzbetriebe bekannt.¹¹ Entsprechend wird als Grund für die Nichtteilnahme am Förderprogramm von 48 Prozent der Referenzbetriebe die Unwissenheit über die Möglichkeit der geförderten Mobilitätsberatung angeführt (vgl. Abb. 5). Bei den Referenzbetrieben sehen 30 Prozent in den Auslandspraktika von Auszubildenden keinen Nutzen für das Unternehmen; 27 Prozent äußern, kein Interesse an Auslandspraktika für Auszubildende zu haben. Für 15 Prozent sprechen ein hoher Verwaltungsaufwand und mangelnde Kapazitäten gegen ein Auslandspraktikum für Auszubildende. Mit einer verstärkten Beratung der Betriebe könnte diesen Hemmnissen entgegengewirkt werden. Zwölf Prozent der Referenzbetriebe haben kein Interesse an Praktika von Auszubildenden aus dem Ausland in ihrem Unternehmen, und zehn Prozent der Referenzbetriebe bezweifeln den Lerneffekt für Auszubildende.¹² Die Bereitschaft zur Entsendung besteht allerdings bei 44 Prozent der befragten Referenzbetriebe.¹³ Dies verweist auf ein hohes Potenzial von neuen Betrieben, die durch die Mobilitätsberatung angesprochen werden könnten.

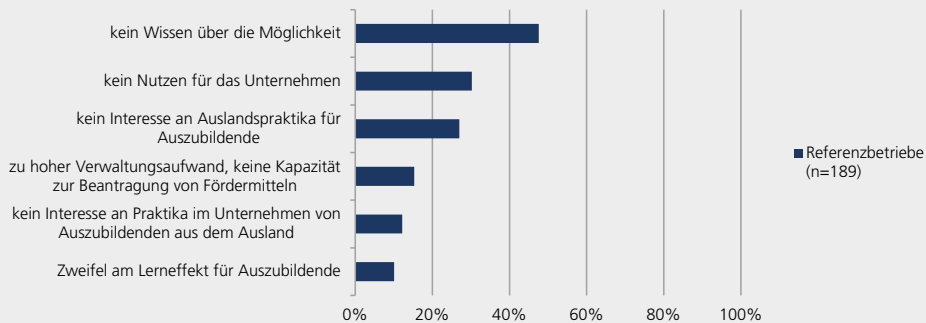
11 Insgesamt haben 186 Referenzbetriebe auf diese Frage geantwortet.

12 Insgesamt haben 189 Referenzbetriebe auf diese Frage geantwortet.

13 Insgesamt haben 161 Referenzbetriebe auf diese Frage geantwortet.

Abbildung 5: Gründe für die Nichtteilnahme am Förderprogramm

Warum nimmt Ihr Unternehmen nicht an der Beratung im Förderprogramm "Berufsbildung ohne Grenzen" teil?



Quelle: eigene Darstellung

Die Experteninterviews mit den Mobilitätsberatern und -beraterinnen spiegeln Unsicherheit bezüglich der Aufgabenstellung der betrieblichen Mobilitätsberatung wider: Die Mobilitätsberater/-innen haben verschiedene Schwerpunktsetzungen, und auch beim Selbstverständnis gibt es verschiedene Auffassungen. Es gibt Mobilitätsberater/-innen, die sich ausschließlich als Berater/-innen verstehen, und andere, die auch bei der kompletten Umsetzung der Auslandspraktika stark beteiligt sind. Viele Mobilitätsberater/-innen wünschen sich eine klare Definition der Beratung und der damit verbundenen Arbeitsweisen. Die Aufgaben, die in den Förderrichtlinien definiert sind, könnten angepasst bzw. weiter ausgeführt werden.

Unsicherheiten im Vorgehen zeigen sich auch bei der Ansprache neuer Unternehmen. In den Experteninterviews mit den Mobilitätsberatern und -beraterinnen hat sich deutlich herausgestellt, dass eine Kaltakquise von Unternehmen mit einigen Herausforderungen verbunden ist und daher nur von wenigen regelmäßig betrieben wird. Nicht jedem Mobilitätsberater bzw. jeder Mobilitätsberaterin ist es möglich, auf Unternehmenskontakte der eigenen Kammer zurückzugreifen und diese für die Akquise zu nutzen. Eine indirekte Ansprache der Betriebe geschieht oftmals über Berufsschulen und Auszubildende. Einige Mobilitätsberater/-innen versenden Informationsmaterial direkt an die Auszubildenden, um Teilnehmende zu gewinnen. Allerdings zeigen die Interviews mit Beratern und Beraterinnen, dass auf das Versenden von Informationsmaterial via E-Mail oder über die Post nur wenig Rückmeldung komme. Ein Nachfassen via Telefon sei meist ebenfalls wenig erfolgreich. Eine Kaltakquise auf Messen mit einer gezielten Ansprache von Unternehmen und Auszubildenden wird von einigen der befragten Mobilitätsberater/-innen als deutlich erfolgreicher angesehen. Andere Formen bilden die Information an Kreishandwerkerschaften, die Teilnahme an Veranstaltungen (z. B. Ausbildertreffen, Innungsversammlungen) und Vorträge im Rahmen von überbe-

trieblichen Ausbildungen. Einige Kammern unterstützen ihre Mobilitätsberater/-innen bei der Akquise, indem sie z. B. bei der Lehrstellenbörse die Möglichkeit eingefügt haben, ein Interesse an Mobilitätsberatung anzugeben.

Viele der befragten Mobilitätsberater/-innen berichteten aber auch, dass es eine Veränderung bei der Beratung und Akquise gibt. So werden die Berater/-innen zunehmend von Unternehmen und Auszubildenden bezüglich des Themas Mobilitätsberatung angefragt. Die positiven Erfahrungen anderer Unternehmen und Auszubildenden in der Region verbreiten sich und überzeugen vom Mehrwert.

Aus den Experteninterviews mit den Mobilitätsberatern und -beraterinnen wird deutlich, wie unterschiedlich die Akquise der einzelnen Berater/-innen ist. Die Entwicklung einer strukturierten, systematischen Vorgehensweise und Strategie bei der Betriebsakquise wäre empfehlenswert. Bestehende Akquisewege und Unterstützungsmöglichkeiten der Kammern und die Überzeugungsarbeit bei Betrieben könnten verstärkt und weiterentwickelt werden. Die Kooperation mit Berufsschulen stellt einen wichtigen Multiplikatoreffekt bei der Akquise von Betrieben dar.

Um die Auslandsmobilität bei Betrieben zu verankern, sollte die strategische Beratung der Betriebe ausgebaut werden. Ziele, Inhalte, Argumente, Ansatzpunkte, Wege der Ansprache etc. sind deutlicher zu definieren. Nur so kann den Mobilitätsberatern und -beraterinnen mehr Sicherheit und eine strukturierte Vorgehensweise gegeben werden. Mobilitätsberater/-innen sollten in der gezielten Ansprache der Betriebe, bei der Gewinnung und Überzeugung neuer Betriebe durch Arbeitshilfen und Qualifizierung unterstützt werden. Es ist notwendig, eine Kommunikationsstrategie und Öffentlichkeitsarbeit hinsichtlich der Kernzielgruppe der KMU auszubauen, um das Marktsegment und Alleinstellungsmerkmal der betrieblichen Mobilitätsberatung offensiver zu besetzen und damit neue Betriebe auf das Angebot der Mobilitätsberatung aufmerksam zu machen.

4.3 Internationale Aktivitäten der Kammern

Die teilnehmenden Kammern sehen sehr viele Vorteile in einer Förderung durch das Programm. Genannt wurden hier die Steigerung der Auslandsmobilität der teilnehmenden Betriebe (83 %), die Steigerung internationaler Kooperationen der teilnehmenden Betriebe (63 %), die Steigerung der Internationalität der teilnehmenden Betriebe (60 %) und der Zuwachs von ausländischen Partnern im Kammerbezirk (33 %).¹⁴

Die Kammern, die am Förderprogramm teilnehmen, zeigen ein langfristiges Interesse an Mobilität: 89 Prozent geben an, bereits vor 2015 am Förderprogramm teilgenommen zu haben.¹⁵

Grundsätzlich besitzt die Mobilitätsberatung einen hohen Stellenwert für die Auslandsmobilität. Das Ergebnis der Befragung der teilnehmenden Betriebe zeigt, dass 78 Prozent

14 Insgesamt haben 30 Kammern, die am Förderprogramm teilnehmen, auf diese Frage geantwortet.

15 Insgesamt haben 28 Kammern, die am Förderprogramm teilnehmen, auf diese Frage geantwortet.

ohne eine Mobilitätsberatung keine Entsendung durchgeführt hätten.¹⁶ Auch der Umstand, dass die Betriebe nach einer Entsendung zu 85 Prozent angeben, erneut Entsendungen vornehmen zu wollen, verdeutlicht dies.¹⁷ Die Unterstützung durch die Kammern wird von den befragten Mobilitätsberatern und -beraterinnen sehr unterschiedlich wahrgenommen. Die Freiräume, die den Beratern und Beraterinnen gegeben werden, sind kammerabhängig und auch der Stellenwert der Mobilitätsberatung innerhalb der Kammern ist sehr verschieden. Bei den meisten Kammern stellt die Beratung laut den Mobilitätsberatern und -beraterinnen eher ein „Nischenprodukt“ dar, und die Einbindung in andere Abteilungen findet kaum statt. Nur bei wenigen Kammern ist die Auslandsmobilität eine wichtige Strategie, die umfassend unterstützt wird. Dies widerspricht den zuvor genannten positiven Einschätzungen zur Auslandsmobilität und Internationalität.

In der Onlinebefragung der Kammern ohne Teilnahme am Förderprogramm BoG wurde als Grund für die Nichtteilnahme besonders häufig angegeben, dass Auslandsmobilität keine Priorität im Kammerbezirk hat (47 %). Deutlich weniger Kammern benennen fehlende Kapazitäten für eine Antragstellung (19 %), kein Interesse der Unternehmen an Auslandsmobilität (11 %), kein Bedarf (9 %), keine Annahme von Fördermitteln (7 %) und vorhandene Mobilitätsberatung an anderer Stelle (7 %).¹⁸ 78 Prozent der Kammern ohne Beteiligung am Förderprogramm äußern, international aktiv zu sein, und 80,4 Prozent zeigen einen Bedarf ihrer Mitgliedsbetriebe an Auslandspraktika an.¹⁹ Hier wird ein Widerspruch zwischen Bedarf an Mobilitätsberatung und der Priorität des Themas Auslandsmobilität deutlich. Trotz internationaler Aktivitäten der Kammern wird der Zusammenhang zur Auslandsmobilität und dem Bedarf der Betriebe an Auslandspraktika anscheinend nicht gesehen. 50 Prozent der Kammern mit einer Mobilitätsberatung, die nicht im Förderprogramm sind, wünschen sich laut Befragung Unterstützung bei der Vermittlung an bestehende Mobilitätsprojekte, und 20 Prozent von ihnen wünschen sich eine Übernahme der Antragstellung. Erkennbar wird, dass mehr Kammern Mobilitätsprogramme nutzen, ohne dass sie eine Mobilitätsberatung unterhalten. Es besteht demnach ein hohes Potenzial, Kammern als neue Projektteilnehmer im Förderprogramm zu gewinnen.

Wachsende Bedeutung von Internationalität und internationale Aktivitäten werden durch die befragten (teilnehmenden und nichtteilnehmenden) Kammern bestätigt. Dies weist auf ein Potenzial hin, das für eine Internationalisierung der Kammern aufgegriffen werden könnte. Im Förderprogramm BoG gibt es Vorreiter bei den Kammern, die vernetzte internationale Aktivitäten mit Mobilitätsberatung und Auslandsmobilität etabliert haben und praktizieren. Aus Sicht der Evaluation sollten sowohl den Kammern als auch den Unternehmen Argumente für die Förderung von Mobilität geliefert werden. Aber auch Kammern

16 Insgesamt haben 207 teilnehmende Betriebe auf diese Frage geantwortet.

17 Insgesamt haben 211 teilnehmende Betriebe auf diese Frage geantwortet.

18 Insgesamt haben 57 Kammern, ohne Teilnahme am Förderprogramm, auf diese Frage geantwortet.

19 Insgesamt haben 50 bzw. 51 Kammern, die nicht am Förderprogramm teilnehmen, auf diese Frage geantwortet.

außerhalb des Programms sollten die Vorteile einer verbesserten Vernetzung internationaler Aktivitäten und der Mehrwert von Auslandspraktika nähergebracht werden.

5 Schlussfolgerungen zur Wirkungsforschung in der Berufsbildung

Bei dieser Evaluation handelte es sich um eine Erfolgskontrolle nach § 7 Abs. 2 BHO, die dazu eingesetzt wurde, Handlungsempfehlungen für das nächste Förderprogramm und die Entwicklung einer neuen Förderrichtlinie von BoG zu gewinnen. Anhand der dargestellten Ergebnisse wird der Nutzen der Auslandsmobilität und der betrieblichen Mobilitätsberatung von Betrieben und Auszubildenden bestätigt. Bei der Akquise neuer Betriebe für die Auslandsmobilität von Auszubildenden und der Internationalisierung der Kammern konnten Handlungsbedarfe identifiziert werden, die im Rahmen der neuen Programmförderlinie verstärkt werden könnten.

Im vorliegenden Beitrag wurden nur ausgewählte Ergebnisse dargestellt, jedoch konnten im Rahmen der Fokusgruppendifkussion mit den Akteuren viele weitere Handlungsempfehlungen für eine neue Förderrichtlinie rückgekoppelt und bestätigt werden (vgl. MOHORIČ u. a. 2018). Die Durchführung und Ergebnisse dieser Evaluation haben gezeigt, dass die drei Komponenten der Erfolgskontrolle – Zielerreichungs-, Wirkungs- und Wirtschaftlichkeitskontrolle – gut geeignet sind, um Handlungsempfehlungen zur verbesserten Förderprogrammgestaltung zu erhalten. Insofern könnte das Instrument Erfolgskontrolle nach § 7 Abs. 2 BHO häufiger im Rahmen von Programmevaluationen in der Berufsbildung angewendet werden, da es zur Weiterentwicklung von Förderlinien und -schwerpunkten dienen kann.

Zur Verbesserung bzw. Erweiterung des verwendeten Methodensets und zur Vertiefung der Erkenntnisse sind folgende alternativen Methoden denkbar:

Qualitative Interviews oder eine Gruppendiskussion mit ausgewählten Kammern hätten die internationalen Aktivitäten der Kammern intensiver in den Fokus nehmen können. Hierbei hätte genauer untersucht werden können, wie internationale Aktivitäten der Kammern und Auslandsmobilität oftmals nebeneinanderstehen und innerhalb der Kammern wenig verknüpft werden. Diese Ergebnisse hätten die Gründe dafür näher beleuchtet und reflektiert. Dies könnte mehr Verständnis zur Verbindung aller internationaler Aktivitäten bei den Kammern fördern und Ansatzpunkte für Verbesserungen liefern.

Für das Förderprogramm und die dazugehörigen Mobilitätsberater/-innen wäre eine Sammlung von Best-Practice-Fallbeispielen aus Beratung und Akquise hilfreich. Dadurch würden erfolgreiche Akquisestrategien und Beratungsweisen unterstützt, vereinheitlicht und intern verbreitet. Dies würde nicht nur die Einarbeitung neuer Mitarbeitender fördern, sondern auch den langjährigen Beratern und Beraterinnen neue Ansätze aufzeigen. Dies hätte in Gruppendiskussionen mit Mobilitätsberatern und -beraterinnen entwickelt werden können.

Um der in den Experteninterviews deutlich gewordenen Unklarheit gegenüber der eigenen Aufgabenstellung und dem Vorgehen der Mobilitätsberater/-innen zu begegnen, erge-

ben sich zwei mögliche Ansatzpunkte für weiteren Forschungsbedarf hinsichtlich der Beratungsqualität.

Die notwendigen Definitionen, Fallbeispiele und der Bedarf an transparenten Beratungsstrukturen und Vorgehensweisen sprechen in der Gesamtbetrachtung bei Beratungen für die Einführung einer qualitativen Kennzahl. Denn bislang werden nur quantitative Kennzahlen erhoben: die Anzahl der durchgeführten Beratungen und Informationsveranstaltungen, die Anzahl der beratenen/teilnehmenden Betriebe und die durchgeführten Auslandspraktika. Eine qualitative Kennzahl für die Beratung wird von den befragten Mobilitätsberatern und -beraterinnen und den Teilnehmenden der Fokusgruppe befürwortet. Die Entwicklung eines qualitativen Indikators bzw. die darauf basierende Erfassung von Kennzahlen wird jedoch als Herausforderung gesehen und erfordert deshalb einen gemeinsamen Prozess, so die Auffassung der Fokusgruppe. In thematischen Gruppendiskussionen zu Beratungsbereichen mit Mobilitätsberatern und -beraterinnen und der Koordinierungsstelle könnten verschiedene Vorgehensweisen aufgezeigt und kategorisiert werden, um Qualitätskriterien der Mobilitätsberatung zu entwickeln. Ein solcher Ansatzpunkt war jedoch im Evaluationsauftrag nicht enthalten und konnte daher nicht verfolgt werden.

Die durchgeführte Evaluation bewegte sich im Rahmen der Strukturen, die das Förderprogramm BoG den Akteuren und Handlungsfeldern vorgibt. Wenn Auslandspraktika und Mobilität ein wichtiger Attraktivitätsfaktor für Berufsausbildung sind, dann wird das Instrument der (betrieblichen) Mobilitätsberatung immer relevanter und sollte weiter erforscht werden. Folgende Fragen für eine Untersuchung des Instruments der Mobilitätsberatung könnten daher lauten: Was beinhaltet Mobilitätsberatung? Wie stellt sich der dahinterstehende Beratungsprozess dar? Was bedeutet Qualität erfolgreicher Mobilitätsberatungen? Interessant wäre es, alle existierenden Mobilitätsprogramme in den Fokus zu nehmen, um die dort vollzogenen Beratungen von folgenden Mobilitätsberatergruppen zu untersuchen:

- ▶ Mobilitätsberater/-innen bestehender bilateraler Programme, z. B. mit Frankreich und Polen, oder auch grenzüberschreitender Mobilitätsprogramme wie z. B. INTERREG (Regionalprogramm der Europäischen Union);
- ▶ Mobilitätsberater/-innen ohne BoG-Förderung, die bei den Kammern und anderen Trägern angesiedelt sind;
- ▶ Mobilitätsberater/-innen der Informations- und Beratungsstelle für Auslandsaufenthalte in der beruflichen Bildung (IBS bei der NA beim BIBB);
- ▶ Mobilitätsberater/-innen der NA beim BIBB, die vorwiegend antragstellende Betriebe im Programm Erasmus+ beraten;
- ▶ Berufsschullehrer/-innen, die Mobilitätsberatungen an berufsbildenden Schulen durchführen;

- Mobilitätsberater/-innen im Programm „Integration durch Austausch“ der ESF-Richtlinie Bund, das sich speziell an Personen mit besonderen Schwierigkeiten beim Zugang zu Arbeit und Ausbildung, z. B. benachteiligte Zielgruppen (u. a. Menschen mit Behinderung), richtet.

Die Fragen, was Mobilitätsberatung im Einzelnen beinhaltet und wie der Beratungsprozess gestaltet ist, könnten in qualitativen Interviews und Gruppengesprächen mit ausgewählten Mobilitätsberatern und -beraterinnen aller Programme erörtert werden.

Literatur

- BOHNSACK, Ralf: Gruppendiskussion. In: FLICK, Uwe; VON KARDORFF, ERNST; STEINKE, Ines (Hrsg.): Qualitative Forschung. Ein Handbuch. 7. Aufl. Reinbek 2009, S. 369–384
- BUNDESAMT FÜR WIRTSCHAFT UND AUSFUHRKONTROLLE (BAFA) (Hrsg.): Richtlinie zur betrieblichen Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften („Berufsbildung ohne Grenzen“). 2015b – URL: https://www.bundesanzeiger.de/ebanzwww/wexsservlet?page.navid=official_starttoofficial_view_publication&session.sessionid=fb6df9af6185d9c0d9c33914e4ee8919&fts_search_list.selected=cf3001f5a518bb58&&fts_search_list.destHistoryId=03231&fundstelle=BAAnz_AT_28.12.2015_B1 (Zugriff: 31.05.2018)
- BUNDESINSTITUT FÜR BERUFSBILDUNG (BIBB) (Hrsg.): Datenreport zum Berufsbildungsbericht 2017. Bielefeld 2017 – URL: https://www.bibb.de/datenreport/de/datenreport_2017.php (Zugriff: 09.03.2021)
- BUNDESINSTITUT FÜR BERUFSBILDUNG (BIBB) (Hrsg.): Datenreport zum Berufsbildungsbericht 2018. Bonn 2018 – URL: https://www.bibb.de/dokumente/pdf/bibb_datenreport_2018.pdf (Zugriff: 29.07.2020)
- BUSSE, Gerd; FROMMBERGER, Dietmar: Internationalisierung der Berufsausbildung als Herausforderung für Politik und Betriebe. Düsseldorf 2016
- FRIEDRICH, Werner; KÖRBEI, Markus: Verdeckte Mobilität in der beruflichen Bildung – Ermittlung von Auslandsaufenthalten in der Erstausbildung außerhalb des EU-Programms für lebenslanges Lernen und der bilateralen Austauschprogramme des Bundesministeriums für Bildung und Forschung. Bonn 2011
- KOMMISSION DER EUROPÄISCHEN GEMEINSCHAFTEN (Hrsg.): EMPFEHLUNG DER KOMMISSION vom 6. Mai 2003 betreffend die Definition der Kleinstunternehmen sowie der kleinen und mittleren Unternehmen (2003/361/EG). In: Amtsblatt der Europäischen Union L 124 vom 20.05.2003, S. 36–41
- MOHORIČ, Andrea u. a.: Evaluation des BMWi-Förderprogramms zur betrieblichen Beratung zur Erhöhung der grenzüberschreitenden Mobilität von Auszubildenden und jungen Fachkräften („Berufsbildung ohne Grenzen“ – BoG). Abschlussbericht. Berlin 2018 –

URL: https://www.f-bb.de/fileadmin/PDFs-Publikationen/Evaluation_BoG_Endbericht_f-bb_11.2018_gekuerzt_zur_Veroeffentlichung_final.pdf (Zugriff: 30.09.2019)

NATIONALE AGENTUR BILDUNG FÜR EUROPA BEIM BIBB (NA BEIM BIBB) (Hrsg.): Mobilität Chancen nutzen – Erklärung Mobilitätspass. 2018 – URL: <https://www.europass-info.de/dokumente/mobilitaet/> (Zugriff: 31.08.2018)

NATIONALE AGENTUR BILDUNG FÜR EUROPA BEIM BIBB (NA BEIM BIBB) (Hrsg.): Auslandsaufenthalte in der Berufsausbildung 2017. Mobilitätsstudie. 2018 – URL: https://www.na-bibb.de/fileadmin/user_upload/na-bibb.de/Dokumente/06_Metavigation/02_Service/Publikationen_Warenkorb/Studien_impuls/NA_Mobilitaetsstudie_WebV3_180706-2.pdf (Zugriff: 09.03.2021)

► Über die Autorinnen und Autoren

Prof. Dr. Marc Beutner

Universität Paderborn

Prof. Dr. Ernst Deuer

Duale Hochschule Baden-Württemberg Ravensburg

Prof. Dr. Hubert Ertl

Bundesinstitut für Berufsbildung

Peter Jablonka

SALSS – Sozialwissenschaftliche Forschungsgruppe

Birte Komosin

Arbeit und Leben – DGB/VHS Berlin-Brandenburg

Nadja Konrad

Forschungsinstitut Betriebliche Bildung (f-bb)

Henning Kruse

Arbeit und Leben – DGB/VHS Berlin-Brandenburg

Dr. Junmin Li

Universität zu Köln

Adj. Prof. Dr. Wolfgang Meyer

Centrum für Evaluation (CEval)

Andrea Mohorič

Forschungsinstitut Betriebliche Bildung (f-bb)

Dr. Rosemarie Sackmann

Forschungsinstitut Betriebliche Bildung (f-bb)

Dr. Lydia Schulze Heuling

Professorin für Physikdidaktik an der Western Norway University of Applied Sciences in Bergen, Norwegen

Prof. Dr. Reinhard Stockmann

Centrum für Evaluation (CEval)

Dr. Philipp Ulmer

Bundesinstitut für Berufsbildung

Dr. Steffen Wild

Duale Hochschule Baden-Württemberg Ravensburg

Dr. Wolfgang Wittig

Forschungsinstitut Betriebliche Bildung (f-bb)

► Zur Arbeitsgemeinschaft Berufsbildungsforschungsnetz

Die Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN) ist ein freiwilliger Zusammenschluss von Einrichtungen, die Beiträge zur Berufsbildungsforschung aus unterschiedlichen wissenschaftlichen Disziplinen leisten. Ziel der Arbeitsgemeinschaft ist es, die wissenschaftliche Zusammenarbeit zu verbessern, den Austausch von Forschungsergebnissen, Meinungen und Erfahrungen zu unterstützen, relevante Forschungsfelder zu identifizieren und den wissenschaftlichen Nachwuchs zu fördern.

Die AG BFN wurde am 7. Juni 1991 in Nürnberg gegründet. Gründungsmitglieder sind die Sektion für Berufs- und Wirtschaftspädagogik (BWP) der Deutschen Gesellschaft für Erziehungswissenschaften (DGfE), das Bundesinstitut für Berufsbildung (BIBB) und das Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB). Über die Jahre hinweg hat sich der Kreis der Netzwerkpartner erweitert. Neben der Sektion BWP, dem IAB und dem BIBB haben sich die pädagogischen Institute der Bundesländer sowie zahlreiche Forschungsinstitute in privater und öffentlicher Trägerschaft dem Netzwerk angeschlossen. Eine Kooperation mit weiteren wissenschaftlichen Gruppierungen und Gesellschaften, die Berufsbildungsforschung betreiben, wird angestrebt. Die Netzwerkpartner verpflichten sich zu folgenden Arbeitsgrundsätzen: Sie

- leisten Beiträge zur grundlagen- oder anwendungsorientierten Berufsbildungsforschung;
- unterziehen ihre Forschungsarbeiten einer kontinuierlichen Qualitätsentwicklung,
- veröffentlichen die Ergebnisse ihrer Arbeit,
- beteiligen sich am wissenschaftlichen Diskurs,
- fördern den wissenschaftlichen Nachwuchs und
- beteiligen sich an den Aktivitäten der AG BFN.

In ihren Aufgaben wird die AG BFN vom BIBB gefördert und unterstützt. Sie führt themenorientierte Foren, Fachtagungen und Workshops durch. Die Veranstaltungen der AG BFN werden in der Regel auf Basis eines Call for Papers durchgeführt. Die eingereichten Beiträge werden begutachtet und von den Organisatoren der Veranstaltung in Abstimmung mit dem Vorstand der AG BFN ausgewählt. Seit 2006 werden die über den Call for Papers ein-

gegangenen und in dieser Schriftenreihe veröffentlichen Beiträge durch fachlich ausgewiesene Wissenschaftler/-innen des Netzwerks begutachtet. Im zweijährigen Turnus wird der Friedrich-Edding-Preis für Berufsbildungsforschung vergeben – eine Auszeichnung für den wissenschaftlichen Nachwuchs. Alle Aktivitäten der AG BFN werden auch im Internetportal www.agbfn.de dokumentiert. Zudem bietet das Portal einen Bereich für Ankündigung und einen Zugang zum VET Repository des BIBB, das aus der Literaturlatenbank Berufliche Bildung, einem Gemeinschaftsprojekt der AG BFN, hervorgegangen ist.

In dieser Schriftenreihe erschienene Veröffentlichungen:


- Berufsbildung 4.0. Steht die berufliche Bildung vor einem Umbruch? Birgit ZIEGLER; Ralf TENBERG [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 26. Bonn: Bundesinstitut für Berufsbildung, 2020
- Zugang zu beruflicher Bildung für Zuwandernde. Britta MATTHES; Eckart SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 25. Bonn: Bundesinstitut für Berufsbildung, 2020
- Berufliche Integration durch Sprache. Jörg ROCHE; Thomas HOCHLEITNER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 24. 2. Auflage. Bonn: Bundesinstitut für Berufsbildung, 2020
- Multidisziplinär – praxisorientiert – evidenzbasiert: Berufsbildungsforschung im Kontext unterschiedlicher Anforderungen. Reinhold WEISS; Eckart SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 23. Bonn: Bundesinstitut für Berufsbildung, 2018
- Didaktik der beruflichen Bildung – Selbstverständnis, Zukunftsperspektiven und Innovationsschwerpunkte. Tade TRAMM; Marc CASPER; Tobias SCHLÖMER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 22. Bielefeld: Bertelsmann, 2018
- Berufsbildung für Geringqualifizierte. Barrieren und Erträge. Britta MATTHES; ECKART SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 21. Bielefeld: Bertelsmann, 2017
- Entwicklungen und Perspektiven in den Gesundheitsberufen – aktuelle Handlungs- und Forschungsfelder. Ulrike WEYLAND; Karin REIBER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 20. Bielefeld: Bertelsmann, 2017
- Verzahnung beruflicher und akademischer Bildung. Duale Studiengänge in Theorie und Praxis. Uwe FASSHAUER; Eckart SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 19. Bielefeld: Bertelsmann, 2016
- Inklusion in der Berufsbildung: Befunde – Konzepte – Diskussionen. Andrea ZOYKE; Kirsten VOLLMER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 18. Bielefeld: Bertelsmann, 2016
- Wirkung von Fördermaßnahmen im Übergangssystem – Forschungsstand, Kritik, Desiderata. Heike SOLGA; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 17. Bielefeld: Bertelsmann, 2015

- Sicherung des Fachkräftepotenzials durch Nachqualifizierung. Befunde – Konzepte – Forschungsbedarf. Eckart SEVERING; Martin BAETHGE [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 16. Bielefeld: Bertelsmann, 2015
- Individuelle Förderung in heterogenen Gruppen in der Berufsausbildung. Befunde – Konzepte – Forschungsbedarf. Eckart SEVERING; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 15. Bielefeld: Bertelsmann, 2014
- Weiterentwicklung von Berufen – Herausforderungen für die Berufsbildungsforschung Eckart SEVERING; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 14. Bielefeld: Bertelsmann, 2014
- Akademisierung der Berufswelt? Eckart SEVERING; Ulrich TEICHLER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 13. Bielefeld: Bertelsmann, 2013
- Qualitätsentwicklung in der Berufsbildungsforschung. Eckart SEVERING; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 12. Bielefeld: Bertelsmann, 2012
- Berufliches Bildungspersonal – Forschungsfragen und Qualifizierungskonzepte. Philipp ULMER; Reinhold WEISS; Arnulf ZÖLLER [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 11. Bielefeld: Bertelsmann, 2012
- Prüfungen und Zertifizierungen in der beruflichen Bildung. Anforderungen – Instrumente – Forschungsbedarf. Eckart SEVERING; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 10. Bielefeld: Bertelsmann, 2011
- Migration als Chance. Ein Beitrag der beruflichen Bildung. Mona GRANATO; Dieter MÜNK; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 9. Bielefeld: Bertelsmann, 2011
- Kompetenzermittlung für die Berufsbildung. Verfahren, Probleme und Perspektiven im nationalen, europäischen und internationalen Raum. Dieter MÜNK; Andreas SCHELTEN [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 8. Bielefeld: Bertelsmann, 2010
- Theorie und Praxis der Kompetenzfeststellung im Betrieb – Status quo und Entwicklungsbedarf. Schriften zur Berufsbildungsforschung der Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN). Dieter MÜNK; Eckart SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 7. Bielefeld: Bertelsmann, 2009
- Qualität in der beruflichen Bildung. Forschungsergebnisse und Desiderata. Hans-Dieter MÜNK; Reinhold WEISS [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 6. Bielefeld: Bertelsmann, 2009
- Zukunft der dualen Berufsausbildung – Wettbewerb der Bildungsgänge. Schriften zur Berufsbildungsforschung der Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN). Hans DIETRICH; Eckart SEVERING [Hrsg.]. Berichte zur beruflichen Bildung AG BFN, 5. Bielefeld: Bertelsmann, 2008
- Perspektiven der Berufsbildungsforschung: Orientierungsleistungen der Forschung für die Praxis. Ergebnisse des AG BFN-Expertenworkshops vom 15. und 16. März 2006 im Rahmen der Hochschultage Berufliche Bildung in Bremen. Reinhold NICKOLAUS; Arnulf ZÖLLER [Hrsg.]. AG BFN, 4. Bielefeld: Bertelsmann, 2007

- Der europäische Berufsbildungsraum – Beiträge der Berufsbildungsforschung. 6. Forum der Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN) 19.–20. September 2005, Universität Erfurt. Manfred ECKERT; Arnulf ZÖLLER [Hrsg.]. AG BFN, 3. Bielefeld: Bertelsmann, 2006
- Vollzeitschulische Berufsausbildung – eine gleichwertige Partnerin des dualen Systems. Arnulf ZÖLLER [Hrsg.]; Manfred KREMER [Mitarb.]; Günter WALDEN [Mitarb.]; Dieter EULER [Mitarb.]; u. a. Berichte zur beruflichen Bildung AG BFN, 2. Bielefeld: Bertelsmann, 2006
- Klassifizierungssystem der beruflichen Bildung. Entwicklung, Umsetzung und Erläuterungen. Franz SCHAFFEL-KAISER. Bielefeld: Bertelsmann, 2005

► Abstract

What does science contribute to vocational education and training (VET) in practice? How should projects, programmes, and political strategies be further developed? The aim of evaluation, design research, and impact research is to analyse conditions for the development of projects, programmes, pilot projects, or system reforms. Implementation processes are accompanied systematically and in a research-led manner. Effects of interventions are identified and attributed to their causes. The VET research network (AG BFN) dedicates this volume to evaluation and impact research in VET. New trends, concepts, and methods are addressed.



Was kann die Wissenschaft zur Berufsbildungspraxis beitragen? Wie sollen Projekte, Programme und politische Strategien weiterentwickelt werden? Ziel von Evaluation, Begleit- und Wirkungsforschung ist, Bedingungen für die Entwicklung von Projekten, Programmen, Modellvorhaben oder Systemreformen zu analysieren. Implementationsprozesse werden wissenschaftlich begleitet. Wirkungen von Interventionsmaßnahmen sollen erfasst und ihren kausalen Ursachenfaktoren zugeschrieben werden. Mit diesem Band widmet sich die Arbeitsgemeinschaft Berufsbildungsforschungsnetz (AG BFN) der Evaluations- und Wirkungsforschung in der beruflichen Bildung. Es werden neue Trends, Konzepte und Methoden thematisiert.

Bundesinstitut für Berufsbildung
Robert-Schuman-Platz 3
53175 Bonn

Telefon (0228) 107-0

Internet: www.bibb.de
E-Mail: zentrale@bibb.de



ISBN 978-3-8474-2941-8