

Aus dem Institut für Europäische Sportentwicklung und Freizeitforschung

der Deutschen Sporthochschule Köln

Geschäftsführender Leiter: Prof. Dr. Jürgen Mittag

**Zwischen Verwissenschaftlichung und
Anwendungsorientierung:**

**Evaluationsforschung und Akteursstrukturen im Kontext
von „Sport für Entwicklung“**

an der Deutschen Sporthochschule Köln

zur Erlangung des akademischen Grades

Doktorin der Sportwissenschaft

angenommene Dissertation

vorgelegt von

Katrin Bauer

aus

Reutlingen

Köln 2022

Erster Gutachter: Prof. Dr. Jürgen Mittag
Zweite Gutachterin: Prof. Dr. Katja Siefken
Vorsitzende*r des Promotionsausschusses: Prof. Dr. Mario Thevis
Datum der Disputation: 14. September 2022

Eidesstattliche Versicherungen gem. § 7 Abs. 2 Nr. 9 der Promotionsordnung der Deutschen Sporthochschule Köln, 30.03.2020:

Hierdurch versichere ich:

Ich habe diese Arbeit selbständig und nur unter Benutzung der angegebenen Quellen und technischen Hilfen angefertigt; sie hat noch keiner anderen Stelle zur Prüfung vorgelegen. Wörtlich übernommene Textstellen, auch Einzelsätze oder Teile davon, sind als Zitate kenntlich gemacht worden.

Hierdurch erkläre ich, dass ich die „Leitlinien guter wissenschaftlicher Praxis“ der Deutschen Sporthochschule Köln eingehalten habe.

Köln, den 1. Dezember 2022



Datum, Unterschrift

Vorwort

Die vorliegende Arbeit stellt eine leicht überarbeitete Version für die Veröffentlichung dar. Wenige Quellenangaben, Fehlerstellen sowie Tabellen wurden behutsam korrigiert oder aktualisiert.

Ich bedanke mich bei Mitarbeitenden des Kooperationspartners der Drittmittelprojekte, der Gesellschaft für Internationale Zusammenarbeit und Entwicklung (GIZ), die den Zugang zu Informationen und Quellen ermöglicht oder erleichtert haben. Darüber hinaus bedanke ich mich bei allen Kolleg*innen des Instituts für Europäische Sportentwicklung und Freizeitforschung für die jahrelange Unterstützung.

Kurzfassung

Die vorliegende Arbeit leistet einen Beitrag (1) zur Verortung des Forschungsfeldes ‚Sport für Entwicklung‘ (SfE), (2) zur Positionierung von Forschenden im Themenfeld und (3) zur Dokumentation der deutschen sportbezogenen Maßnahmen in der Entwicklungszusammenarbeit (EZ).

Die Dokumentation der sportbezogenen Maßnahmen in Deutschland verdeutlicht, dass ‚der Sport‘ in der deutschen EZ zwar früh Einzug gehalten hat, dies international jedoch aufgrund bislang fehlender Dokumentationen kaum bekannt ist. Erstmals werden die deutschen Bemühungen ab ihren Anfängen in den 1970er Jahren bis heute, inklusive inhaltlicher Schwerpunkte und Akteurskonstellationen, abgebildet. Die Betrachtung der SfE-Maßnahmen der *Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)* vor dem Hintergrund pädagogischer und sportwissenschaftlicher Ansätze und Erklärungsmodelle zeigt, dass v.a. inhaltliche und funktionslogische Überschneidungen zur Erlebnispädagogik existieren. Zudem wird deutlich, dass die Zusammenarbeit mit Akteur*innen aus dem Sport maßgeblichen Einfluss auf die inhaltliche Ausgestaltung hat. Aus einer terminologischen Reflexion resultiert, dass unklar bleibt, wie viele entwicklungs-spezifische Inhalte Teil der SfE-Maßnahmen sind und ob eher eine *Entwicklung des Sports* (Sportförderung) nach dem *Sport-für-Alle*-Prinzip und durch den Ausbau von Infrastruktur stattfindet.

Zur Untersuchung von Evaluationsaktivitäten und Akteursstrukturen im allgemeinen EZ-Kontext, im Kontext SfE sowie im Kontext der wissenschaftlichen Begleitung des Sektorvorhabens SfE der GIZ als evaluationspraktisches Beispiel wird eine Perspektive aus der Evaluationsforschung gewählt. Durch die Betrachtung der drei Kontexte mittels aufgestellter Untersuchungskriterien zu den Evaluationsaktivitäten (Definitionen, Ansätze, Funktionen, Qualität) und den Akteursstrukturen (Bestimmung von beteiligten Stakeholdern, Rollen und Rollenverteilungen) wird deutlich, wie ‚wissenschaftlich‘ bzw. ‚anwendungsorientiert‘ das Forschungsfeld gestaltet und wie dies durch beteiligte Akteur*innen beeinflusst ist. Das genutzte Modell arbeitet u.a. definitorische Schwächen, einen verwirrenden Theoretisierungsprozess (‚jede*r nutzt was er*sie kennt‘), einen

Mangel an Evaluationen zur Qualitätssicherung (Meta-Evaluationen) und ex-post-Evaluationen sowie Herausforderungen im Hinblick auf die Evaluation von Wirkungen und Nachhaltigkeit und damit verbundene Einflussfaktoren (Organisation/intern, Umwelt/extern) heraus. Auch Problematiken bzgl. unterschiedlicher oder einseitiger Funktionszuschreibungen sowie auf eingebundene Stakeholder werden ersichtlich. Der Fokus auf die Qualitätssicherung von Evaluationen und die damit einhergehende Nutzung der Evaluationsstandards verdeutlichen Missstände im allgemeinen und sportbezogenen EZ-Kontext, v.a. auch, was die Partizipation von Beteiligten und die Unabhängigkeit von Evaluierenden anbelangt. Durch die Betrachtung aus der Perspektive der Evaluationsforschung wird zwar deutlich, dass im SfE-Kontext weit mehr Evaluationsforschung betrieben als deklariert wird, allerdings stößt das Modell auch an Grenzen, bspw. sind klare Typisierungen von Studien nicht immer möglich.

Resultierend aus den Ergebnissen kommt die Autorin zu dem Schluss, dass sich das Forschungsfeld aus dem Spannungsverhältnis zwischen Verwissenschaftlichung und Anwendungsorientierung nur dann lösen kann, wenn Veränderungen sowohl auf Seiten der Forschenden als auch auf Seite der (staatlichen) EZ-Akteur*innen stattfinden. Eine Weiterentwicklung ist möglich, wenn es gelingt, sich *von einem multidisziplinären zu einem interdisziplinären Forschungsfeld* zu entwickeln. Um dies zu erreichen, müssen sich SfE-Forschende aus ihrer disziplinären Isolation lösen und damit anders positionieren: Sie müssen eine gemeinsame Sprache (v.a. bzgl. Definitionen) und gemeinsame Problemsichten (Orientierung an den SDGs) finden, und durch den Blick in andere Forschungs- und Evaluationsbereiche (v.a. in die allgemeine EZ) theoretische Ansätze und Methoden vereinheitlichen. Im Hinblick auf die EZ-Akteur*innen bedeutet dies, dass Evaluation sich wünschenswerterweise freier und flexibler gestalten ließe und sich damit *aus der ‚Zwangsjacke der Auftraggeberwünsche‘ befreit*. Schlussendlich kann das Forschungsfeld jedoch nur dann Bestand haben, wenn sich *stabilere, nachhaltige Finanzierungsstrukturen* entwickeln, die bislang weder über das EZ- noch über das Hochschulsystem gegeben sind. Optimistisch kann in beiden Fällen mit der Nachhaltigkeitsagenda argumentiert werden, die auf beide Systeme auch künftig einen großen Einfluss haben wird und i.R. derer Zielerreichung der Sport eine umfangreiche Rolle spielen kann.

Abstract

This thesis contributes (1) to the localization of the research field ‚Sport for Development‘ (SfD), (2) to the positioning of researchers in the field and (3) to the documentation of German sport-related measures in development cooperation.

The documentation of sport-related measures in Germany illustrates that although ‚sport‘ was already part of German development cooperation at an early stage, this is hardly known internationally due to a lack of documentation so far. For the first time, the German efforts from their beginnings in the 1970s until today, including content priorities and actor constellations, are mapped. By examining the SfD measures of the GIZ against the background of pedagogical and sports science approaches and explanatory models, it becomes clear that there are overlaps with experiential education, especially in terms of content and functional logic. In addition, it shows that partner sport organisations had a significant influence on the content of the programme activities. A terminological reflection shows that it remains unclear how much development-specific content is part of the SfD activities and whether it is more the *development of sport* according to ideas embedded in *sport-for-all* principle and through the physical development of infrastructure.

To examine evaluation activities and actor structures in the context of general development cooperation, in the context of SfD, and in the context of the scientific monitoring of the GIZ's SfD sector programme as a practical example of evaluation, a perspective from evaluation research is chosen. By examining the three contexts with the help of established criteria for evaluation activities (definitions, approaches, functions, quality) and stakeholder structures (determination of stakeholders involved, roles and distribution of roles), it becomes clear how ‚scientific‘ or ‚applied‘ the research field is and how this is influenced by the stakeholders involved. This model highlights definitional weaknesses, a confusing theorization process ('everyone uses what they know'), that evaluation is often reduced to a one-sided accountability function, a lack of evaluations for quality assurance (meta-evaluations) and ex-post evaluations, as well as challenges with regard to the evaluation of impacts and sustainability and related influencing factors (organization/internal, environment/external). The focus on the quality assurance of

evaluations and the associated use of evaluation standards reveals shortcomings in the general and sports-related context of development cooperation, especially with regard to the participation of stakeholders and the independence of evaluators. By looking from the perspective of evaluation research, it becomes clear that much more evaluation research is conducted in the SfD context than declared, but the model also has its limits, e.g. clear typifications of studies are not always possible.

Based on the results, the author concludes that the research field can only escape from tensions between science and applied orientations if changes take place both on the part of the researchers and on the part of the (governmental) actors in development cooperation. Further development is possible if the research field succeeds in *evolving from multidisciplinary to interdisciplinarity*. To manage this, SfD researchers must break free from their disciplinary isolation and thus position themselves differently. They need to find a common language (especially regarding definitions) and common problem views (orientation towards the SDGs), and to unify theoretical approaches and methods by looking into other research and evaluation fields (especially into general development cooperation). With regards to actors in development cooperation, this means that evaluation processes should be freer and more flexible and thus *be freed from the 'straitjacket of the client's wishes'*. Ultimately, however, the field of research can only endure if more *stable, sustainable financing structures* develop, which neither the system in development cooperation system nor the university system provide so far. Optimistically, in both cases, it can be argued that the sustainability agenda will continue to have a major impact on both systems and that sport can play a major role in achieving its goals.

Inhaltsverzeichnis

Abkürzungsverzeichnis	i
Abbildungsverzeichnis	iv
Tabellenverzeichnis	iv
1. Einleitung: (Evaluations-)Forschung im Kontext ‚Sport für Entwicklung‘ als Notwendigkeit und Herausforderung	1
1.1 Ausgangspunkt und Relevanz	1
1.2 Forschungsstand und -desiderate	4
1.3 Zielsetzungen der Arbeit.....	7
1.4 Aufbau der Arbeit	9
2. Hintergrund: ‚Sport für Entwicklung‘ in Deutschland	12
2.1 Anfänge der internationalen Sportförderung und ‚verzögerte Blütezeit‘ von ‚Sport für Entwicklung‘	12
2.2 Neupositionierung von Sport in der dt. EZ: Sport für Entwicklung ab 2013 ...	22
2.2.1 Vorhaben und Projekte der GIZ.....	22
2.2.2 Kooperationsstrukturen und deutsche SfE-Akteur*innen	34
2.3 Terminologische Reflexion: Der Versuch einer begriffsbestimmenden Einordnung von ‚Sport für Entwicklung‘ bzw. von Sport im deutschen EZ-Kontext ..	43
2.3.1 Von der Sportförderung zu ‚Sport für Entwicklung‘?	43
2.3.2 Von welchem ‚Sport‘ reden wir überhaupt?.....	47
2.3.3 Was ist der ‚deutsche‘ SfE-Ansatz? Präzisierung der sportbezogenen EZ- Maßnahmen	54
3. Rahmung: Theoretisch-konzeptionelle Untersuchungsperspektiven und methodische Vorgehensweisen	59
3.1 Evaluation	59
3.1.1 Definitionen	59
3.1.2 Evaluationsansätze und -klassifikation.....	63
3.1.3 Zur Funktion bzw. zum Nutzen von Evaluation	70
3.1.4 Zur Qualität von Evaluationen.....	74
3.2 Evaluationsgegenstände und -schwerpunkte.....	77
3.2.1 Definitionen: Projekte, Programme, Vorhaben und Maßnahmen	77
3.2.2 Zum Verlauf von Maßnahmen.....	79

3.2.3	Evaluation von Wirkungen und Nachhaltigkeit von Maßnahmen	80
3.3	Akteur*innen im Evaluationsprozess	86
3.3.1	Rollenverteilung und institutionelle Zugehörigkeit.....	86
3.3.2	Beteiligte Stakeholdergruppen.....	87
3.3.3	Zur Rolle von Evaluierenden.....	89
3.4	Untersuchungsperspektiven.....	93
3.5	Methodische Vorgehensweise: Literatur- und Dokumentenrecherche.....	97
4.	Betrachtung des EZ-Kontexts: Entwicklung von Evaluation in Phasen	101
4.1	Phase I: Die Anfänge (1960er und 1970er).....	103
4.2	Phase II: Rasante Entwicklungen und erste Krisen (1980er)	106
4.3	Phase III: Gezielte Umgestaltungen und komplexe Verhältnisse (1990er) ...	113
4.4	Phase IV: Konsolidierung und Institutionalisierung (2000-2015).....	121
4.5	Phase V: Alte und neue Herausforderungen (seit 2015).....	137
4.6	Evaluation in der EZ – ein Zwischenfazit.....	148
5.	Betrachtung des SfE-Kontexts: Evaluation und Forschung	153
5.1	Definitionen und Funktionsverständnis von Evaluation.....	153
5.2	Evaluierende und Forschende aus SfE-Perspektive.....	156
5.3	Erste Entwicklungen eines Forschungsfeldes	159
5.3.1	Das Forschungsfeld SfE – ein Überblick.....	159
5.3.2	Kontext und Meilensteine	162
5.3.3	Institutionalisierungs- und Professionalisierungsprozesse	165
5.3.4	Standardisierungsprozesse.....	169
5.4	Schwerpunkte, theoretische Zugänge und Ansätze	171
5.4.1	Inhaltliche Schwerpunkte	171
5.4.2	Zur Fundierung und ‚Theoretisierung‘ von SfE.....	173
5.4.3	Genutzte Ansätze im SfE-Kontext.....	179
5.4.4	Zur Wirkungsdebatte im SfE-Kontext.....	185
5.5	SfE-Evaluation und -Forschung in Deutschland	189

6. Evaluation in der Praxis: Kooperation und Begleitforschung im Rahmen des Sektorvorhabens ‚Sport für Entwicklung‘ der GIZ	197
6.1 Überblick: Aufgaben, Arbeitsbereiche u. Beteiligte im Evaluationsprozess .	197
6.2 Phase I: Orientierung	200
6.3 Phase II: Strukturierung, Standardisierung und Scaling-up	205
6.4 Phase III: Capacity Development und Wirkungsmessung	213
6.5 Zusammenfassung und Bewertung der Evaluationsprozesse	215
7. Diskussion: Zwischen Verwissenschaftlichung und Anwendungsorientierung .	220
7.1 Das Forschungsfeld SfE und SfE-Forschende	220
7.1.1 Entwicklung: Parallelen zur EZ-Evaluation – zeitlich verzögert und wenig genutzt	220
7.1.2 Definitionen: Diskursstiftende Ungenauigkeiten	223
7.1.3 Theoretische Fundierung, Ansätze und Methoden: alle nutzen, was sie kennen – wenige nutzen, was vorhanden ist.....	224
7.1.4 Funktion und Nutzen: Evaluation ist mehr als ‚accountability‘ und ‚learning‘.....	231
7.1.5 Evaluation von Wirkungen und Nachhaltigkeit von Maßnahmen: Herausfordernd, komplex und unterfinanziert.....	233
7.1.6 Qualität: zu wenig Meta-Evaluationen, nicht-genutzte Standards.....	238
7.1.7 Zur Unabhängigkeit von beauftragten, externen Evaluierenden.....	240
7.1.8 Zur Partizipation von (zu) vielen Beteiligten	243
7.2 Schlussfolgerungen und Ausblicke	246
7.3 Reflexion und Perspektiven	254
7.3.1 Möglichkeiten und Grenzen	254
7.3.2 Theoretisch-konzeptionelle Perspektiven.....	256
8. Fazit und Ausblick: Interdisziplinarität und Gestaltungsspielräume als wegweisende Voraussetzungen	262
Literaturverzeichnis	271

Abkürzungsverzeichnis

3iE	International Initiative for Impact Evaluation
AA	Auswärtiges Amt
BMI	Bundesministerium des Inneren
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
bspw.	beispielsweise
bzgl.	bezüglich
ca.	circa
CD	Capacity Development
CEval	Centrum für Evaluation der Universität des Saarlandes
d.h.	das heißt
DAC	Development Assistance Committee (Ausschuss für Entwicklungshilfe der <i>OECD</i>)
DED	Deutscher Entwicklungsdienst
DeGEval	Gesellschaft für Evaluation (Deutschland und Österreich)
DEval	Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit
DEZA	Schweizer Direktion für Entwicklung und Zusammenarbeit
DFB	Deutscher Fußballbund
DIE	Deutsches Institut für Entwicklungspolitik
DLV	Deutscher Leichtathletik Verband
DOSB	Deutscher Olympischer Sportbund
DSB	Deutscher Sportbund
DSHS	Deutsche Sporthochschule Köln
dt. EZ	Deutsche Entwicklungszusammenarbeit
ebd.	ebenda
et al.	et alii bzw. et aliae
etc.	et cetera
EZ	Entwicklungszusammenarbeit
FGD	Focus Group Discussions (Fokusgruppendifkussion)
FIFA	Fédération Internationale de Football Association (Internationaler Fußballverband)
ggf.	gegebenenfalls
GIZ	Gesellschaft für internationale Zusammenarbeit

GTZ	Gesellschaft für technische Zusammenarbeit (<i>GIZ</i> ab 2010)
i.R.	im Rahmen
ICSSPE	International Council of Sport Science and Physical Education (Weltrat für Sportwissenschaft und Leibes-/Körpererziehung)
InWEnt	Internationale Weiterbildung und Entwicklung gemeinnützige GmbH
IOC	International Olympic Committee (Internationales Olympisches Komitee)
KfW	Kreditanstalt für Wiederaufbau
M&E	Monitoring und Evaluation
MDG(s)	Millennium Development Goal(s)
MINEPS	International Conference of Ministers and Senior Officials responsible for Physical Education and Sport (Internationale Sportministerkonferenz)
Mio.	Millionen
n.b.	nicht bekannt
NGO	Non-Governmental Organisation (Nichtregierungsorganisation)
NOK	Nationales Olympisches Komitee
NONIE	Network of Networks on Impact Evaluation
o.D.	ohne Datum
OECD	Organization of Economic Cooperation and Development (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung)
PE	Physical Education
PM	Performance Management
PRA	Participatory Rural Appraisal
RBM	Results Based Management
RCT	Randomised Control Trial
RV	Regionalvorhaben
RV S4DA	Regionalvorhaben „Sport für Entwicklung in Afrika“ der GIZ
S.	Seite
s.	siehe
s.o.	siehe oben
SDG(s)	Sustainable Development Goal(s) (Nachhaltige Entwicklungsziele)
SDP	Sport for Development and Peace
SDP IWG	Sport for Development and Peace International Working Group (der UN)
SFD, S4D	Sport for Development
SfE	Sport für Entwicklung

sfw	streetfootballworld
sog.	sogenannte
SV SfE	Sektorvorhaben „Sport für Entwicklung“ der GIZ
ToR	Terms of Reference
u.a.	unter anderem
u.a.	unter anderem
UN	United Nations, Vereinte Nationen
UN GA	United Nations General Assembly
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children’s Fund
UNOSDP	United Nation Office on Sport for Development and Peace
USAID	United States Agency for International Development
usw.	und so weiter
v.a.	vor allem
vgl.	vergleiche
WM	Weltmeisterschaft
WoM	Wirkungsorientiertes Monitoring
YDF	Youth Development through Football
z.B.	zum Beispiel
z.T.	zum Teil
ZOPP	Zielorientierte Projektplanung (theoriegeleiteter Ansatz der GTZ)

Abbildungsverzeichnis

Abb. 1: Fördermittel internationale Sportförderung AA und BMZ im Zeitverlauf	21
Abb. 2: Kooperationsstrukturen der GIZ-Vorhaben in den Partnerländern	34
Abb. 3: Deutsche SfE-Stakeholder	42
Abb. 4: Sportförderung und Sport für Entwicklung	47
Abb. 5: Bandbreite deutscher SfE-Maßnahmen.....	58
Abb. 6: Logisches Modell/Logical Framework	67
Abb. 7: Funktionen/Nutzen von Evaluationen	73
Abb. 8: Lebensverlaufsmodell einer Maßnahme	80
Abb. 9: Logisches Modell mit Zielebenen von Wirkungsevaluationen	81
Abb. 10: Theoretisch-konzeptionelle Rahmung der Arbeit.....	96
Abb. 11: Aufgaben, Arbeitsbereiche im Prozess der wissenschaftlichen Begleitung ..	197

Tabellenverzeichnis

Tab. 1: Fördermittel der Bundesregierung für die Internationale Sportförderung des AA und BMZ 1964-1999	16
Tab. 2: Fördermittel der Bundesregierung für die Internationale Sportförderung des AA und BMZ 2000-2017	20
Tab. 3: Übersicht der Partnerländer der SfE-Vorhaben der GIZ	27
Tab. 4: Standards für Evaluation	74
Tab. 5: Bibliographische Zugänge.....	99
Tab. 6: Scaling-up Strategien in den Partnerländern des SV SfE	207
Tab. 7: Durchgeführte Evaluationen Phase I-III.....	215
Tab. 8: Struktur des Evaluationsleitfadens zum CEval-Ansatz	258
Tab. 9: Checkliste für Evaluierende zur Überprüfung von Evaluationsstandards im Evaluationsprozess	260

1. Einleitung: (Evaluations-)Forschung im Kontext ‚Sport für Entwicklung‘ als Notwendigkeit und Herausforderung

1.1 Ausgangspunkt und Relevanz

Entwicklungsminister Gerd Müller forderte am Ende seiner Amtszeit mit der Vorlage des 16. Entwicklungsberichts der Bundesregierung mehr Geld für Entwicklungszusammenarbeit (Koch, 2021, 20. Oktober). Doch anders als im 14. und 15. Entwicklungsbericht, findet ‚der Sport‘ hier *keine* Erwähnung mehr (vgl. BMZ, Mai 2013; März 2017; September 2021).¹

Dadurch stellt sich die unvermeidbare Frage, ob ‚der Sport‘ künftig in der deutschen Entwicklungszusammenarbeit (dt. EZ) Bestand haben wird, nachdem das *Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ)* dem Ansatz *Sport für Entwicklung (SfE)* in der letzten Dekade besondere Aufmerksamkeit beigemessen hat. Durch die Einrichtung des *Sektorvorhabens „Sport für Entwicklung“ (SV SfE)* in der *Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ)* im Jahr 2013 und der Etablierung weiterer regionaler Vorhaben in Afrika und Nah-Ost, ist SfE zu einem übergreifenden Thema mit zahlreichen Verbindungen zu verschiedenen entwicklungspolitischen Zielen der Bundesregierung avanciert (BMZ, 2010-2020a). Die GIZ als Durchführungsorganisation setzt(e) seit Beginn über 50 sportbezogene Maßnahmen in 40 verschiedenen Ländern um. Nach gut neun Jahren Projektimplementierung, der Renovierung und des Neubaus von ca. 150 Sportplätzen, der Ausbildung von rund 8200 Lehrpersonen, Trainer*innen und Sozialpädagog*innen² zu sog. *SfE-Coaches* und/oder *SfE-Instructors*³

¹ In *kursiv* werden jene Begriffe gesetzt, die hervorgehoben werden sollen. In doppelten Anführungszeichen („“) stehen direkte Zitate, einfache Anführungszeichen (,) werden genutzt, wenn die Autorin einen spezifischen Begriff hervorheben möchte und/oder sich von dem Begriff distanziert.

² In dieser Arbeit werden möglichst *geschlechtsneutrale Formulierungen* bei der Bezeichnung von Personengruppen verwendet. Wenn dies nicht möglich ist, werden *Sternchen (*)* genutzt. Bei feststehenden, zusammengesetzten Bezeichnungen (bspw. Partnerländer) wird nicht gegendert. Zitate werden im Original beibehalten.

³ Als *SfE-Coach* wird eine Person bezeichnet, die Training oder Sportunterricht für Kinder und Jugendliche mit SfE-Schwerpunkt anleitet und durchführt. Ein Coach kann ein*e Trainer*in eines Vereins, Lehrkraft, Sozialarbeiter*in sein oder einen anderen Beruf ausüben. Als *SfE-Instructor* wird eine Person bezeichnet, die andere Coaches in SfE-Ansätzen und Methoden weiterbildet und damit in der Erwachsenenbildung tätig ist (Petry & Bauer, 2019). In der Arbeit werden im Folgenden die englischen, geschlechtsneutralen Begriffe verwendet.

(S4D Resource Toolkit, o.D.) und damit der Ausgabe von Steuergeldern in Millionenbeträgen, stellt sich zunehmend die *Frage nach der Wirksamkeit* dieser Projekte, v.a. auch hinsichtlich ihrer Nachhaltigkeit und nach der spezifischen Wirkung des Ansatzes SfE.

Generell zweifelt die Bevölkerung an der Wirksamkeit der dt. EZ (Schneider & Gleser, 2018). Im letzten Meinungsmonitor zur Entwicklungspolitik zeigt sich, dass besonders „Informationen zur Wirksamkeit von EZ-Projekten [...] dazu beitragen [können], dass die Bevölkerung solche Projekte positiver bewertet“ (Schneider, Eger & Sassenshagen, 2021, S. viii). Vor allem der Sport, der in der EZ als „low-cost high-impact tool“ (Kay & Dudfield, 2013, S. 5) eingesetzt wird und seitens der Vereinten Nationen zur „Allzweckwaffe‘ zur Erreichung der SDGs“ (Brand, 2020, S. 113) stilisiert wird, ist seit jeher und zunehmend dem Druck ausgesetzt, seine Wirksamkeit unter Beweis zu stellen. Dabei warnen Digel und Fornhoff (1989, S. 9) bereits in den späten 1980ern davor,

„man sollte sich davor hüten, den Sport mit allgemeinen entwicklungspolitischen Erwartungen zu überfrachten [...], da Wirkungen des Sports in andere Bereiche hinein wohl möglich sind, meist jedoch nicht in einem direkten Kausalzusammenhang zum Sport stehen und deshalb kaum kalkulier- und steuerbar sind“.

Um „die Qualität und Wirksamkeit“ (BMZ, 2010-2020a) der EZ im Bereich SfE zu verbessern, aber auch um die SfE-Maßnahmen und damit die Ausgabe von Steuergeldern durch Evaluationen legitimieren zu können (vgl. Döring & Bortz, 2016; Caspari, 2004; Stockmann, 2007b; 2017; Vedung, 2015; Weiss, 1998), arbeitet die GIZ von Beginn an mit wissenschaftlichen Partner*innen. Das SV SfE wird seit Mitte 2013 bis aktuell in insgesamt drei Phasen von der *Deutschen Sporthochschule Köln (DSHS)* wissenschaftlich begleitet. In dieser Zeit befanden und befinden sich die Evaluierenden häufig in einem *Spannungsfeld* wieder, in dem sie auf der einen Seite im Rahmen (i.R.) von Forschungsaktivitäten *wissenschaftlichen Ansprüchen* gerecht werden will und auf der anderen Seite mit einem *stark praxisorientierten Arbeitsfeld* konfrontiert ist, für das die Umsetzung solcher Ansprüche eine Herausforderung darstellt. Dazu kommt, dass über *anwendungsorientierte Evaluationen* hauptsächlich „Handlungswissen für die Praxis“ geschaffen werden soll (Kromrey, 2001, S. 113); d.h. sowohl für die SfE-Maßnahmen an sich, als auch für die strategische Beratung zum Arbeitsfeld SfE allgemein, sollen evidenzbasierte, also auf Grundlage von Evaluationen eruierte, Handlungsempfehlungen gegeben werden. Ausgehend von diesen Erfahrungen soll diese Arbeit dem Spannungsfeld weiter

auf den Grund gehen: Inwiefern ist es generell im SfE-Kontext präsent? Wie können sich beauftragte Evaluierende hier verorten? Welche Maßnahmen können ergriffen werden, die zum Abbau von Spannungen beitragen?

Neben den GIZ-Projekten wurden in den vergangenen Jahren an der DSHS weitere (Dritt-mittel-)Projekte im Bereich SfE realisiert, bspw. verschiedene Weiterbildungskooperati-onen mit Partner*innen wie dem Deutschen Olympischen Sportbund (DOSB) und Enga-gement Global. Zudem verankerte sich das Thema zunehmend in der Lehre, nicht zuletzt durch die Etablierung des *Masterstudiengangs International Sport Development and Po-litics* im Jahr 2016. Diese Aktivitäten sprechen zum einen für den Wunsch zahlreicher Akteur*innen nach externer, sportwissenschaftlicher Begleitung und Beratung, zum an-deren für ein gesteigertes Interesse seitens der Studierenden. Durch diese Aktivitäten hat sich an der Hochschule das ‚*Forschungsfeld SfE*‘ etabliert und weiterentwickelt.

Nach über zehn Jahren des ‚Miterlebens‘, wie sich ein neues Forschungsfeld an der Hochschule etabliert hat, und des Mitgestaltens durch die Mitarbeit an zahlreichen Pro-jekten, ist es an der Zeit, die *Entwicklungen, Inhalte und Praktiken dieses Forschungsfel-des*, das nach wie vor um Anerkennung im Hochschulkontext kämpft (vgl. Massey et al., 2019), aber auch die eigene *Rolle als Forschende* in diesem Kontext zu reflektieren. Da-bei geht es in erster Linie darum, *wie* im Forschungsfeld gearbeitet wird, d.h. welche theoretischen Ansätze und methodischen Vorgehensweisen genutzt werden. Durch die starke Verknüpfung der Forschungsprojekte der DSHS mit Akteur*innen aus der Politik und dem Sport interessiert zweitens, *wer* an der Gestaltung des Forschungsfeldes maß-geblich beteiligt ist, d.h. welche Akteur*innen, inklusive der Forschenden, es prägen. Diese Ausführungen führen wiederum zum Einstieg dieser Arbeit und lassen weiterfüh-rende Fragen aufkommen: Heißt keine Erwähnung im Sportbericht der Bundesregie- rung, dass die dt. EZ dem Sport den Rücken kehrt? Welche Auswirkungen hat dies mög- licherweise auf das Forschungsfeld SfE in Deutschland? Wie abhängig ist das Forschungs- feld SfE allgemein von EZ-Akteur*innen?

1.2 Forschungsstand und -desiderate

Wie die vorherigen Ausführungen zum Druck, die Wirkungen von EZ-Maßnahmen allgemein und der SfE-Maßnahmen im Speziellen nachzuweisen, verdeutlichen, wird „[...] zwar viel über die Notwendigkeit evidenzbasierter Programme [geredet], aber wenig über die *Notwendigkeit evidenzbasierter Evaluation*“, d.h., es muss ein „evaluative[r] Blick auf die Theorie und Praxis der Evaluation selbst“ geworfen werden (Astbury, 2017, S. 240), worauf diese Arbeit u.a. abzielen soll.⁴ Neben der Betrachtung von theoretischen Ansätzen und der Praxis in der sportbezogenen EZ-Evaluation und Forschung soll gerade die *allgemeine EZ-Evaluation* in den Betrachtungsfokus gerückt werden. Dies liegt zum einen darin begründet, dass SfE-Maßnahmen vermehrt in entwicklungspolitischen Kontexten umgesetzt werden (Giulianotti, Coalter, Collison & Darnell, 2019) und zum anderen, dass der EZ-Evaluation bei der Entwicklung von Evaluationsforschung im Allgemeinen eine besondere, prägende Rolle zugesprochen wird (vgl. Kjekshus, 1991). Hier wurde im Vergleich zu anderen Politikfeldern bereits früh und besonders umfangreich evaluiert sowie darüber diskutiert. Ansätze, Methoden sowie Diskurse um Transparenz und Feedback haben sich „[...] kontinuierlich entlang der Lernkurven des Ressorts professionalisiert und internationalisiert“ (Mayring, Kerlen, Sheikh, Hense & Schwab, 2017, S. 17). Der Blick in die langjährige Evaluationsentwicklung und die damit verbundenen Diskurse in der allgemeinen EZ können möglicherweise richtungsweisend für die sportbezogene EZ-Evaluation sein.

Darüber hinaus existiert generell *wenig Forschung über Evaluation und deren zentrale Probleme* (Meyer & Stockmann, 2017). Gerade „[...] die wissenschaftliche Publikation über Evaluationsprojekte und ihre Verfahren, Probleme und Ergebnisse [ist] grundsätzlich wünschenswert, [da] sie [...] den Fortschritt der Evaluation und die Verbreitung von Wissen im jeweiligen Gegenstandsbereich [fördert]“ (DeGEval, 2016, S. 49), was im Falle dieser Arbeit der SfE-Kontext ist. Zu den zentralen Problemen gehören u.a. die angesprochenen Herausforderungen im beschriebenen Spannungsfeld, die beauftragte Evaluierende zu bewältigen haben. Evaluationsforschung und damit auch Evaluierende be-

⁴ Hervorhebung durch die Autorin.

finden sich in einem *dauerhaften, strukturellen Spannungsverhältnis zwischen Wissenschaft und Praxis*: Während die Praxis auf die Klärung von ‚Wie‘-Fragen abzielt (z.B. Wie kann ein bestimmter Zustand verändert werden? Wie kann eine Zielgruppe erreicht werden?), beschäftigt sich die Wissenschaft mit ‚Warum‘-Fragen (bspw. Warum werden bestimmte Zielgruppen nicht erreicht?), da sie an generellen Zusammenhängen und Erklärungen interessiert ist und nicht an Handlungsempfehlungen, wie sie auch im Zuge von EZ-Evaluationen häufig ausgesprochen werden (Stockmann, 2007).

Dass im EZ-Kontext *externe Evaluierende* oder Gutachtende mit sog. Fremdevaluierungen beauftragt werden, wodurch die Unabhängigkeit von Studien garantiert werden soll, war bereits früh der Fall und ist heutzutage gängige Praxis (Borrmann & Stockmann, 2009b; Caspari, 2004; Reuber, 2006); so auch im Kontext von SfE, wie es die Beauftragung der DSHS durch die GIZ demonstriert. Auch wenn bekannt ist, dass beauftragte Evaluierende aufgrund ihrer verschiedenen Rollen und Handlungsanforderungen zahlreichen Spannungsfeldern ausgesetzt sind, wie den wissenschaftlichen Ansprüchen und der Praxis sowie den verschiedenen Interessen von Beteiligten zeitgleich gerecht zu werden (Bortz & Döring, 2016; Brandt, 2007; Gutknecht-Gmeiner, 2014; Stockmann & Meyer, 2017b), wird die *Forschung zur Rolle von Evaluierenden* und auch die *Forschung zur partizipativen Konzepten*, also die Beteiligung verschiedener Interessengruppen am Evaluationsprozess, als *defizitär* bezeichnet (Gutknecht-Gmeiner, 2014; Heinze, Reiter, Berg-Lupper, Wach & Riedle, 2019, S. 280; Meyer, Funk & Nguyen, 2017). Deswegen sollen diese Aspekte Gegenstände dieser Arbeit sein.

Einen Überblick darüber zu bekommen, welche *Forschungs- bzw. Evaluationsansätze* in der Praxis, auch im Kontext von SfE, genutzt werden, fällt auf den ersten Blick schwer, da sie sich im Laufe der Zeit „wild vermehrt“ haben (Astbury, 2017, S. 238). Dies ist u.a. darauf zurückzuführen, dass Ansätze zum einen aus Jahren praktischer Erfahrung heraus entstanden sind (Darnell, 2012; Maguire, 2013), wie es auch in der EZ-Evaluation der Fall ist. Zum anderen sind sie mit einer oder mehreren Schlüsselpersonen verbunden (Astbury, 2017). Demnach sind im SfE-Kontext Wissenschaftler*innen unterschiedlichster Disziplinen involviert, die alle die disziplinen-eigenen theoretischen Modelle und Methoden

zur Forschung und Evaluation nutzen (Burnett, 2015). Das aktuellere, einschlägige Sammelwerk *Routledge Handbook of Sport for Development and Peace* (Collison, Darnell, Giulianotti, & Howe, 2019) bietet erstmalig eine ausführliche Übersicht zu unterschiedlichen theoretischen und methodischen Zugängen aus der Sozialforschung zum Themenfeld (ebd., S. 163ff.). Des Weiteren geben Literaturübersichten i.R. von Qualifikationsarbeiten (Biermann, 2016, S. 10ff.), vereinzelte (meta-)analytische Artikel (Welty Peachey, Schulenkorf & Hill, 2019), Buchbeiträge (Gadais, 2019), systematische Reviews (Schulenkorf, Sherry & Rowe, 2016) oder Beiträge basierend auf Reviews (Darnell, Whitley, Camiré, Massey, Blom, Chawansky, Forde & Hayden, 2019; Whitley, Fraser, Dudley, Yarrow & van der Merwe, 2020) Einblicke zum Forschungsstand, zu Forschungslücken sowie zu theoretischen und konzeptionellen Fortschritten. Darüber hinaus reflektieren Autor*innen den Zustand und die Aussichten des Forschungsfeldes und erörtern weitere Zugänge und Ansätze (bspw. Giulianotti et al., 2019; Schulenkorf, 2017).

In diesem Zusammenhang fehlen generell aber v.a. für den SfE-Kontext empirische Überprüfungen, d.h. die Untersuchung der *Operationalisierungsfähigkeit von Ansätzen* in unterschiedlichen Kontexten (vgl. Astbury, 2017; Meyer & Stockmann, 2017). Dies bezieht sich im Hinblick auf EZ-Maßnahmen, seien es jene mit oder ohne Sportbezug, v.a. auch auf *Ansätze zur Evaluation von Wirkungen und Nachhaltigkeit*. Letztere stehen deswegen zunehmend im Fokus, da die Nachhaltigkeitsagenda der Vereinten Nationen zum strategischen Mittelpunkt aller EZ-Aktivitäten geworden ist (vgl. BMZ, 2018). Für die entwicklungspolitische Praxis wurden bereits unterschiedlich komplexe Evaluationsverfahren zur Wirkungsanalyse untersucht und dokumentiert (vgl. Oberndörfer, Hanf & Weiland, 2010). Nachhaltigkeit ist zwar seit geraumer Zeit ein Evaluierungskriterium der *Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD)* (OECD, 1991), eine Studie des *Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DE-val)* (Euler, Noltze & Verspohl, 2018) zeigt jedoch, dass in der EZ-Praxis die Bewertung von Nachhaltigkeit sehr heterogen und unsystematisch erfolgt, und insgesamt zu kurz greift. Dadurch ist es der EZ-Praxis bislang nicht möglich, Nachhaltigkeit im Sinne der Agenda 2030 abzubilden (ebd.). Eigene Evaluationsansätze existieren hierzu bislang nur wenige (vgl. Capsari, 2004).

Ein weiterer wichtiger Aspekt, der sowohl ein Desiderat darstellt als auch die Relevanz dieser Arbeit verdeutlicht, beruht auf der Tatsache, dass, auch wenn im Kontext der dt. EZ bereits seit über dreißig Jahren Sportprojekte weltweit gefördert werden und seit 2013 durch die Etablierung eigener Vorhaben in der GIZ SfE zu einem prominenten Themenbereich aufgestiegen ist (BMZ, 2010-2020a), *die frühen deutschen Bemühungen*, aber auch die aktuellen im internationalen (Forschungs-)Kontext weitgehend *unbekannt* sind (vgl. van Eekeren, ter Horst & Fictorie, 2013, S. 19ff.). Dies liegt vermutlich darin begründet, dass frühe Veröffentlichungen zu den Maßnahmen sowie wissenschaftliche Aufbereitungen und Studien lediglich in deutscher Sprache verfasst wurden (s. Digel & Fornhoff, 1989; Giebenhain, 1990; Trosien, 1987). Die Dokumentation der aktuellen Maßnahmen ist bislang in ihrer Gänze unzureichend und lediglich durch drei deutsche Beiträge in einem Sammelband (Petry, 2020) repräsentiert (Grafarend & Gaerte, 2020; Klingmann, Mialkas & Weinberg, 2020; Schmid & Arnemann, 2020). Außer einer Veröffentlichung, die anhand sog. Factsheets und Broschüren⁵ die Rolle des Sports in der dt. Entwicklungspolitik analysiert (Schreiner, Mayer & Kastrup, 2020), wurden keine weiteren Beiträge auf Englisch bzw. international veröffentlicht. Die Arbeit leistet in einem ersten Schritt einen Beitrag, der defizitären Dokumentation der deutschen Maßnahmen entgegenzuwirken und soll in einem weiteren Schritt als Grundlage für Publikationen in englischer Sprache genutzt werden.

1.3 Zielsetzungen der Arbeit

Die vorliegende Arbeit leistet erstens einen Beitrag zur *Verortung des Forschungsfeldes SfE* und zweitens zur *Positionierung von Wissenschaftler*innen*, die sich diesem Forschungsfeld zugehörig fühlen.

Sie hat zum Ziel, Forschungs- und Evaluationsaktivitäten im SfE-Kontext umfangreich zu erfassen und diese im Hinblick auf die Evaluationsforschung in der allgemeinen EZ einzuordnen. Aus der Perspektive einer wissenschaftlichen Einrichtung, die hauptsächlich

⁵ Das BMZ und die GIZ haben zahlreiche Zusammenfassungen zu den SfE-Maßnahmen in den Partnerländern und einige ‚Hochglanzbroschüren‘ zur Rolle des Sports in der deutschen EZ veröffentlicht, die alle über das *S4D Resource Toolkit* online veröffentlicht wurden (s. *S4D Resource Toolkit*, o.D.). Die Broschüren werden an dieser Stelle so bezeichnet, da sie „[...] vornehmlich die Erfolge der eigenen Arbeit“ veranschaulichen (vgl. Borrmann, Fasbender, Holthus, Gleich, Reichl & Shams, 1999, S. 26).

i.R. von Drittmittelprojekten im SfE-Kontext forschen kann, interessiert zudem, wie die eigenen Evaluationsaktivitäten i.R. der wissenschaftlichen Begleitung des SV SfE der GIZ und die eigene Rolle hierbei verortet werden können. Von daher werden die Evaluationsaktivitäten der DSHS als evaluationspraktisches Beispiel im Forschungsfeld SfE ebenfalls untersucht. Bei der Analyse dieser drei Bereiche (Evaluation in der EZ, Evaluation und Forschung im SfE-Kontext, wissenschaftliche Begleitung des SV SfE der GIZ) werden zwei Untersuchungsschwerpunkte festgelegt und dazugehörige Forschungsfragen formuliert:

(1) *Evaluationsaktivitäten* im jeweiligen Kontext

- Was wird unter Evaluation im jeweiligen Kontext verstanden?
- Wozu wird evaluiert, d.h. welchen Zweck bzw. Nutzen verfolgen die Evaluationen?
- Wie wird evaluiert, d.h. welche Ansätze, Evaluationsarten, theoretischen Fundierungen und methodischen Vorgehensweisen werden genutzt?
- Was wird evaluiert, d.h. welche Untersuchungsgegenstände/Evaluanden werden betrachtet (Projekte, Programme, etc.) und worauf liegen Untersuchungsschwerpunkte (Wirkungen, Nachhaltigkeit)?

(2) *Akteursstrukturen* im jeweiligen Kontext

- Welche Akteur*innen sind an Evaluationsprozessen beteiligt?
- Welche Rollen und Aufgaben haben beteiligte Akteur*innen, insbesondere die Forschenden bzw. Evaluierenden?
- Welche(n) Einfluss, Auswirkungen haben die Akteur*innen auf den Evaluationsprozess und auf das Forschungsfeld?

Die Untersuchung der zwei genannten Schwerpunkte sollen am Ende der Arbeit Erkenntnisse dazu liefern, wie *wissenschaftlich* bzw. *anwendungsorientiert* das Forschungsfeld SfE gestaltet und wie dies durch beteiligte Akteur*innen beeinflusst ist. Auf Basis der Ergebnisse sollen perspektivisch handlungsweisende Aussagen getroffen werden, welche Entwicklungsmöglichkeiten das Forschungsfeld hat bzw. haben kann und welche Rolle hierbei beteiligte Akteur*innen und SfE-Forschende spielen.

Darüber hinaus werden durch den Betrachtungsgegenstand der wissenschaftlichen Begleitung des SV SfE der GIZ durch die DSHS die Entwicklung der deutschen Maßnahmen in der sportbezogenen EZ, einschließlich beteiligter Akteur*innen, ihrer Struktur, Organisation, sowie den inhaltlichen Schwerpunkten und Zielsetzungen, seit 2013 erstmalig in ihrer Gänze und in Verbindung mit ihren Anfängen dokumentiert. Die Arbeit leistet demnach einen Beitrag zur *ganzheitlichen Dokumentation* und zum *Wissensmanagement* des Themas SfE in der dt. EZ.

1.4 Aufbau der Arbeit

Aufgrund der bislang unzureichenden Dokumentation beginnt die Arbeit im zweiten Kapitel mit einer ausführlichen *Bestandsaufnahme von SfE in der dt. EZ*. Es beschreibt retrospektiv die Anfänge der internationalen Sportförderung in Deutschland, die ersten Entwicklungen von SfE in der dt. EZ bis hin zum Aufstieg des Themas durch die Etablierung verschiedener Vorhaben in der GIZ. Um einen Überblick über die zahlreichen Maßnahmen zu gewährleisten, sowie diese erstmalig in ihrer Gänze zu dokumentieren, werden hier die einzelnen Vorhaben inklusive ihrer Länderkomponenten und thematischen Schwerpunkte sowie relevante Akteure und Partner beschrieben. Um die SfE-Maßnahmen der dt. EZ verorten zu können, schließt das Kapitel mit einer begriffsbestimmenden Einordnung der sportbezogenen Aktivitäten der dt. EZ. Hierbei gilt es die Maßnahmen, die sowohl solche zur Sportförderung als auch solche zur Nutzung des Sports zur Erreichung entwicklungspolitischer Zielsetzungen umfassen, vor dem Hintergrund internationaler Entwicklungen und der Sportberichtserstattung der Bundesregierung zu reflektieren. Die Arbeit wäre keine sportwissenschaftliche, würde sie sich nicht die Frage stellen, von welchem ‚Sport‘ im Kontext SfE überhaupt die Rede wäre. Durch sportwissenschaftliche Erklärungsmodelle des Sports sowie der Einordnung i.R. pädagogischer Teildisziplinen, wird deutlich, was SfE-Maßnahmen allgemein umfassen. Vor diesem theoretischen Hintergrund können die aktuellen SfE-Maßnahmen der dt. EZ schlussendlich präzisiert werden.

Kapitel 3 führt in unterschiedliche Kategorien ein, die den *theoretisch-konzeptionellen Analyserahmen* der Arbeit vorgeben: Im ersten Teilkapitel werden Definitionen, Ansätze

und Typen, sowie Funktionsweisen und Nutzungsarten von Evaluation erfasst sowie Evaluationsstandards zur Qualitätssicherung erläutert. Folgendes Teilkapitel nimmt Maßnahmen, d.h. Projekte und Programme, als Evaluationsgegenstände in den Blick und führt in die Wichtigkeit der Evaluation von Wirkungen und Nachhaltigkeit von Maßnahmen ein. Ein weiteres Teilkapitel beschäftigt sich mit den unterschiedlichen Akteur*innen im Evaluationsprozess, wobei neben möglichen Rollenverteilungen vor allem auf die Rolle von Evaluierenden eingegangen wird. Kapitel 3.4 fasst alle Untersuchungsperspektiven zusammen. In Kapitel 3.5 werden methodische Vorgehensweisen zu dieser Arbeit erläutert.

In Kapitel 4, 5 und 6 wird *die Evaluationsforschung in der allgemeinen EZ, Evaluation und Forschung im Kontext SfE* und *die wissenschaftliche Begleitung des SV SfE der GIZ durch die DSHS* als praktisches Beispiel vor dem Hintergrund der theoretisch-konzeptionellen Rahmung betrachtet. Dadurch, dass Evaluationsforschung in der EZ sich entlang entwicklungspolitischer Ereignisse entwickelt hat, wird Kapitel 4 chronologisch abgehandelt und führt durch fünf Phasen mit unterschiedlichen Entwicklungsschwerpunkten. Im Blickpunkt hierbei sind, neben prägenden entwicklungspolitischen Meilensteinen, die Einfluss auf die Evaluationsforschung im entsprechenden Politikfeld hatten, Standardisierungs-, Institutionalisierungs- und Professionalisierungsprozesse, wegweisende Studien, sowie inhaltliche, theoretische und methodische Schwerpunkte von Evaluationen. Nach der Betrachtung von Definitionen und Funktionen von Evaluationen sowie der Rolle von Evaluierenden und Forschenden im SfE-Kontext werden im fünften Kapitel die Entwicklungen des Forschungsfeldes nachgezeichnet, wobei ebenfalls auf Meilensteine sowie Institutionalisierungs-, Professionalisierungs- und Standardisierungsprozesse eingegangen wird. Da das Forschungsfeld aus der Perspektive von Forschenden unterschiedlichster Disziplinen betrachtet wird, ist das Teilkapitel zu Schwerpunkten, Ansätzen und Methoden umfangreicher gestaltet und geht u.a. explizit auf die Wirkungsdebatte ein. Das Kapitel schließt mit der Beleuchtung von Evaluations- und Forschungsaktivitäten in Deutschland. Kapitel 6 nimmt die Praxis in den Blick und betrachtet evaluationsspezifische und begleitende Aufgaben der DSHS i.R. der drei-phasigen wissenschaftlichen Begleitung des SV SfE der GIZ, wobei jede Phase für einen anderen inhaltli-

chen Schwerpunkt steht. In den Teilkapiteln werden besonders die durchgeführten Evaluationen, aber auch damit verbundene Umsetzungsmöglichkeiten und -herausforderungen beschrieben. Das letzte Teilkapitel betrachtet die Evaluationen in ihrer Gesamtheit und sortiert die Evaluationsprozesse anhand von Evaluationsstandards ein.

Vor dem Hintergrund der Betrachtungen in Kapitel 4, 5 und 6 wird im siebten Kapitel zum einen ‚*das Forschungsfeld*‘ SfE und zum anderen ‚*Forschende*‘ im Kontext von SfE im Spannungsfeld zwischen Verwissenschaftlichung und Anwendungsorientierung diskutiert. Der Diskussionsteil endet mit abgeleiteten Schlussfolgerungen und bietet perspektivische Ausblicke und Entwicklungsszenarien. Das Kapitel schließt mit einer Reflexion im Hinblick auf die Möglichkeiten und Grenzen der Arbeit sowie bzgl. der eigenen Rolle als involvierte Evaluierende und gibt perspektivische Ausblicke zur theoretisch-konzeptionellen Rahmung. Die Arbeit endet mit einem Gesamtfazit und Ausblick im achten Kapitel.

2. Hintergrund: ‚Sport für Entwicklung‘ in Deutschland

Im Sinne der Verortung des Arbeits- und Forschungsbereichs dieser Arbeit, soll im folgenden Kapitel eine Einführung der Entwicklungen von SfE in Deutschland gegeben werden. Hierbei wird zuerst auf die Entstehungsgeschichte und die Anfänge des Themas Sport in den Maßnahmen der Bundesregierung eingegangen. Die Verankerung von Sport i.R. eines Sektorvorhabens (SV) der GIZ im Jahr 2013 stellte einen Wendepunkt für das Thema SfE in Deutschland dar. Dementsprechend wird das SV SfE sowie weitere Regionalvorhaben mit ihren Zielsetzungen und Inhalten, inklusive der Maßnahmen in verschiedenen Ländern detailliert vorgestellt. Die anschließende Beschreibung von gängigen Kooperationsstrukturen in den Partnerländern sowie deutschen SfE-Stakeholdern veranschaulicht die Vielzahl und Bandbreite an Akteur*innen, mit denen die deutsche EZ im Themenfeld zusammenarbeitet. Ein separates Kapitel erörtert den Wandel von Begrifflichkeiten rund um die Kombination der Termini *Sport* und *Entwicklung*, um letztendlich eine Antwort auf die Frage zu geben, was ‚den deutschen SfE-Ansatz‘ ausmacht.

2.1 Anfänge der internationalen Sportförderung und ‚verzögerte Blütezeit‘ von ‚Sport für Entwicklung‘

Auch wenn Sport und Bewegung international bereits früh als wichtiges Bildungsinstrument anerkannt war und in Chartas der Vereinten Nationen Erwähnung finden,⁶ wurde Sport in Deutschland lange lediglich als Instrument zum „Ausbau politischer Beziehungen“ genutzt, wenn nicht auch um „die Überlegenheit des gesellschaftlichen Systems“ zu Zeiten des West-Ost Konflikts zu demonstrieren (Giebenhain, 1990, S. 81; Trosien, 1987, S. 29).⁷ Das *Auswärtige Amt (AA)* startete 1961 i.R. der *Auswärtigen Kultur- und Bildungspolitik* seine *internationale Sportförderung*. Anlass waren Anfragen von Sportorganisationen aus Ländern des Globalen Südens an den *Deutschen Sportbund (DSB)* und das *Nationale Olympische Komitee für Deutschland (NOK)*⁸, die sich „Anschluss an

⁶ 1959 in der UN Deklaration für Kinderrechte (UN General Assembly, 1959); 1978 Charta für Leibeserziehungen und Sport der UNESCO (UNESCO, 1978).

⁷ Auf eine ausführliche Darstellung der damaligen Entwicklungspolitik Deutschlands wird an dieser Stelle verzichtet. S. hierzu Giebenhain (1990), der die damalige Sportförderpolitik vor dem Hintergrund des Ost-West- sowie Nord-Süd-Konflikts erläutert (vgl. ebd., S. 74f.).

⁸ Zusammenschluss von DSB und NOK 2006 zum *Deutschen Olympischen Sportbund (DOSB)*.

den internationalen Sportstandard“ versprochen (Digel & Fornoff, 1998, S. 175). Der DSB und das NOK waren erste Kooperationspartner*innen der Bundesregierung im Bereich der Sportförderung, die „Maßnahmen in ausgewählten Wettkampfsportarten“ (ebd., S. 175) förderten. Giebenhain (1990, S. 81) beschreibt diese erste Phase⁹ deutscher Sportförderung wie folgt:

„Entwicklungspolitik wie Sportförderung vollzogen sich unter dem Primat der Außenwirtschafts- und der auswärtigen Kulturpolitik, der deutsche Sport stellt sich als deren Instrument bereitwillig zur Verfügung; die Kooperation von AA und Deutschem Sport hat ihre Grundlage in gemeinsamen außenpolitischen Interessen; Sportförderung war geleitet durch eine modernistische Vorstellung von Entwicklung mit dem Vorbild westlicher Organisationsstrukturen. Sie entsprach damit dem unreflektierten Eurozentrismus der Entwicklungshilfe im Allgemeinen; Die sport- und außenpolitischen Interessen wurden auf der Ebene der Legitimation mit humanitären und sportethischen Motiven überlagert.“

Im Zuge einer Neuorientierung der Entwicklungspolitik, wurden ab 1971/1972 Sportfördermaßnahmen auch durch das BMZ durchgeführt. Diese waren von Anfang an auf die *Förderung des Breitensports* ausgerichtet (vgl. Giebenhain, 1990, S. 85) und wurden als „[...] Maßnahmen zur Bildungshilfe und zur Förderung der sozialen Entwicklung [...]“ mit dem Schwerpunkt auf den Schulsport in den Ländern konzipiert (Digel & Fornoff, 1989, S. 4). Sie verfolgen somit andere Zielsetzungen als die bisherigen Sportfördermaßnahmen und sollen sich so von den Aktivitäten des AA abgrenzen (ebd.). Laut Berichten implizieren diese Maßnahmen bereits ein „ganzheitliches Bildungsverständnis“, das die „Erziehung zum Sport“ sowie die „Erziehung durch Sport“ beinhalten (Trosien, 1987, S. 31).

Allerdings führte die Beteiligung beider Ressorts zu Kompetenzstreitigkeiten und Koordinationsmängeln. Um dem entgegenzuwirken, wurde 1971 ein *Interministerieller Ausschuss (IMA) zur Förderung des Sports in den Entwicklungsländern* gegründet, dem neben den fünf stimmberechtigten Mitgliedern (AA, BMZ, BMI¹⁰, DSB und NOK) auch der *Deutsche Fußball-Bund (DFB)*, als Stellvertretung aller Fachverbände), die *Gesellschaft für*

⁹ Vor der Institutionalisierung der Sportförderung durch das AA, hatten 1933 bereits erste „Sportkontakte“, eingeleitet durch Carl und Liselott Diem in der Türkei, „die als Entwicklungszusammenarbeit im Sport einzustufen sind“, stattgefunden (Giebenhain, 1990, S. 75).

¹⁰ Das Bundesministerium des Inneren (BMI) förderte in erster Linie den deutschen Hochleistungssport. Die Beteiligung des BMI am IMA „resultiert [...] in erster Linie aus dem Sachverhalt, dass dem Bundesminister des Innern die parlamentarische Verantwortung innerhalb der Bundesregierung für die deutsche Sportpolitik und deren Koordinierung zukommt (Digel & Fornoff, 1989, S. 204).

technische Zusammenarbeit (GTZ¹¹, als Hauptdurchführungsorganisation) und das Bundesverwaltungsamt (als Verwaltungsinstitution) angehören.¹² Ab 1974 wurde verstärkt der Versuch unternommen, die unterschiedlichen Förderungsbereiche und Kompetenzen streng voneinander abzugrenzen: Das AA ist für die Förderung des Leistungs- und Verbandssport zuständig, das BMZ für den Schul- und Breitensport sowie für die Sportlehrerausbildung an Hochschulen (Digel & Fornoff, 1989). Giebenhain (1990, S. 93) spricht von „Dauerproblemen [...] trotz dieser formalen Differenzierungen“.¹³ Trotz der Abgrenzungsversuche gelang es dem Ausschuss nicht, die Sportförderung der Bundesregierung konzeptionell zu steuern (Digel & Fornoff, 1989). Giebenhain (1990, S. 102) sieht den Grund in den Abgrenzungsproblemen darin, dass es *keinen theoretischen Bezugsrahmen* gab („konzeptionelles Vakuum“), an dem sich die Diskussion um Funktion, Ziele und Legitimation von „Entwicklungsmaßnahmen im Sport“ hätte orientieren können.

Ein Jahr nach der Gründung der GTZ im Jahr 1974 wurde in der Abteilung für Bildung, Wissenschaft und Kommunikationswesen der *Fachbereich Sport* angesiedelt (Dohrmann, John & Küper, 1985). Dieser führte sowohl für das BMZ als auch für das AA seit Beginn alle Langzeitprojekte durch, die hauptsächlich die Aus- und Fortbildung von Trainer*innen und Lehrkräften in den Sportarten Fußball und Leichtathletik beinhalteten (vgl. Trosien, 1987, S. 74ff.). Er verwaltete somit den größten Teil des Finanzvolumens der sog. „Sportentwicklungshilfe“ (Digel & Fornoff, 1989, S. 179). Der Fachbereich veröffentlichte seit 1976 regelmäßig sog. *Sportreports*, in denen einzelne Maßnahmen der „Sportförderung in Ländern der Dritten Welt“ beschrieben (vgl. Dohrmann et al., 1985, S. 99f.)¹⁴ oder zu bestimmten Themen, bspw. zu *Sportunterricht* oder *Sports for all*, zusammengefasst wurden (vgl. Dohrmann, 1987, 1988). Im Rahmen ihrer Aufträge wollte

¹¹ Später GIZ.

¹² Ausführlicher (Aufgaben der einzelnen Mitglieder, Arbeitsweisen etc.) zu diesem Ausschuss, s. Trosien (1987, S. 117ff.).

¹³ Ein ausschlaggebender Punkt war die Tatsache, dass dieselben Sportexpert*innen (DSB, NOK) und Entwicklungsexpert*innen (GTZ) für beide Auftraggeber arbeiteten. Die GTZ (Dohrmann et al., 1985, S. 24) betonte in ihrer Arbeit die „Gemeinsamkeiten der Auswärtigen Kulturpolitik und der Entwicklungspolitik“. Giebenhain (1990, S. 90ff.) erörtert die Abgrenzungsprobleme auch im Hinblick auf die Sportberichtserstattung ausführlich.

¹⁴ In einer *GTZ Schriftenreihe* wurden abgeschlossene und sich in der Durchführung befindenden Projekte im Bereich der Sportförderung zusammengefasst (vgl. Beyer & Friedel, 1979).

die GTZ „[...] Beiträge zum Aufbau von Sportstrukturen in Entwicklungsländern leisten und diesen damit helfen, sportliche Aktivitäten in breiten Schichten der Bevölkerung zu entfalten und den Anschluss an die internationalen Bemühungen um die Leibeserziehung¹⁵ zu gewinnen“ (Dohrmann et al., 1985, S. 23).¹⁶ Auch wenn die Sportverbände – hauptsächlich über das AA – ebenfalls eingebunden waren, sahen sie in der GTZ einen Konkurrenten in Form einer „neuen Sportbehörde“ (ebd., S. 190), „[die] sich über den Umweg Entwicklungshilfe in das Sportgeschäft gedrängt habe, dabei für derartige Maßnahmen keinen Sachverstand besitze“ (ebd., S. 277).¹⁷

Auch *Hochschulen* waren an den Sportfördermaßnahmen beteiligt, hauptsächlich im Bereich von Sportlehrerfortbildungen und damit häufig in Kooperation mit der GTZ. Vor allem die Universitäten Bayreuth, Darmstadt und Tübingen haben „einen ihrer Schwerpunkte ihrer Lehre und Forschung der Sportentwicklungshilfe gewidmet“ (Digel & Fornoff, 1989, S. 219f.). Die DSHS war damals lediglich durch den Einsatz sog. „Sportexperten“ vertreten, wie bspw. Lieselott Diem für das AA/GTZ/NOK in Brasilien zur Förderung des Breitensports (vgl. Giebenhain, 1990, S. 166f.; Trosien, 1987, S. 124ff.) Der Forschungsstand hinsichtlich der sportbezogenen EZ war laut Digel und Fornoff (1989, S. 349) damals noch „völlig unzureichend“. Eine „begleitende wissenschaftliche Unterstützung“ wurde demnach empfohlen.

Die anfängliche Funktion des Sports in der deutschen Entwicklungspolitik wurde ab Ende der 1970er und Anfang der 1980er Jahre stark kritisiert, u.a. durch den damaligen BMZ-Minister Erhard Eppler, der vor den Gefahren eines „Kulturexports“ warnte, womit die „Zerstörung traditioneller Bewegungskulturen“ einhergeht (Eichberg, 1973; Rittner, 1977 zitiert nach Digel & Fornoff, 1989).¹⁸ Die Kritik löste eine *Debatte über die Legitimation der Sportförderung* sowie über Ziele und Funktionszuschreibungen von Sport in

¹⁵ In der damaligen Terminologie sah die GTZ Schwierigkeiten: Die mit dem Sport assoziierten Begriffe *Körper-* oder *Leibeserziehung* sollten überwunden und Begriffe wie *Bewegung*, *Spiel* und *Sport* genutzt werden (vgl. Dohrmann et al., 1985, S. 32).

¹⁶ Dabei berücksichtigte sie zwei Voraussetzungen: „(1) Sport ist integraler Bestandteil einer umfassenden lebenslangen Bildung der Menschen und daher Bestandteil der Bildungssysteme der Entwicklungsländer“ sowie (2) dass Schul-, Breiten- und Verbandssport zusammengehören und demnach auch Verbindungen in allen Sportfördermaßnahmen gesucht werden sollten (Dohrmann et al., 1985, S. 23).

¹⁷ Eine ausführlichere Beschreibung aus der Perspektive der Verbände gibt Trosien (1987, S. 100ff.).

¹⁸ Die Stellungnahmen des damaligen Ministers werden bei Giebenhain (1990, S. 103ff.) beschrieben, der Begriff des „Kulturexports“ genauer beleuchtet (ebd., S. 217ff.).

der EZ aus, die auch heute noch geführt wird.¹⁹ Digel und Fornoff (1989) fassen i.R. ihres Forschungsvorhabens „Sport in der EZ“, im Auftrag des BMZ, die Funktionszuschreibungen erstmalig zusammen und identifizieren *sechs zentrale Funktionen*: Persönlichkeitsbildung, Integration, Gesundheit, Chancengleichheit und Grundbedürfnisbefriedigung.

Die finanziellen Aufwendungen der Bunderegierungen für die Sportförderung sind in den ersten drei Jahrzehnten insgesamt angewachsen, wobei das Finanzvolumen des AA regelmäßig höher als das des BMZ war. Zwar wurde die Differenz Ende der 1970er und Anfang der 1980er Jahre geringer, allerdings waren die Aufwendungen insgesamt, insbesondere für die Sportförderung des BMZ, rückläufig (Digel & Fornoff, 1989). „Die rückläufige Tendenz ist [...] auf die Systematik der Vergabeverfahren zurückzuführen. Sportprojekte können sich dabei im Wettbewerb mit anderen Sektoren nur selten durchsetzen“ (ebd., S. 180). Während die Summe der Sportförderung des AA auch bis Ende der 1990er Jahren nicht unter 4 Mio. DM fiel, sanken die Ausgaben im BMZ ab Mitte der 1990er kontinuierlich (vgl. Tab. 1; Schott & Merkel, 2009, S. 120f.).

Tab. 1: Fördermittel der Bundesregierung für die Internationale Sportförderung des AA und BMZ 1964-1999 (Dt. Bundestag, 1970, 1973, 1976, 1978, 1982, 1987, 1991, 1995, 1999, 2002)

Jahr	AA ²⁰	BMZ	Jahr	AA	BMZ
1964	500 000 DM	-	1982	5 141 000 DM	4 713 000 DM
1965	610 000 DM	-	1983	4 981 000 DM	4 265 000 DM
1966	685 000 DM	-	1984	5 486 000 DM	3 120 000 DM
1967	710 000 DM	-	1985	4 889 000 DM	2 945 000 DM
1968	785 000 DM	-	1986	5 364 000 DM	3 115 000 DM
1969	913 000 DM	-	1987	5 521 000 DM	2 929 000 DM
1970	1 538 000 DM	-	1988	5 325 000 DM	3 740 000 DM
1971	2 235 000 DM	-	1989	5 806 000 DM	3 602 000 DM
1972	2 815 000 DM	1 650 000 DM	1990	6 002 000 DM	2 794 000 DM
1973	2 581 000 DM	417 000 DM	1991	7 338 000 DM	2 292 000 DM
1974	5 604 000 DM	2 047 000 DM	1992	8 312 000 DM	1 863 000 DM
1975	5 195 000 DM	2 494 000 DM	1993	6 521 000 DM	1 087 000 DM
1976	4 883 000 DM	3 009 000 DM	1994	6 243 000 DM	1 032 000 DM
1977	5 340 000 DM	2 663 000 DM	1995	5 235 000 DM	410 000 DM
1978	5 021 000 DM	2 964 000 DM	1996	4 994 000 DM	179 000 DM
1979	5 673 000 DM	4 280 000 DM	1997	4 492 000 DM	357 000 DM
1980	4 720 000 DM	4 005 000 DM	1998	4 500 000 DM	248 000 DM
1981	5 506 000 DM	4 487 000 DM	1999	5 300 000 DM	512 000 DM

¹⁹ Zur Legitimationsdiskussion und Funktionszuschreibungen aus Sicht der Sportwissenschaft s. Giebenhain (1990, S. 189ff.).

²⁰ Die Zahl bezieht sich lediglich auf die *Sportförderung in den Entwicklungsländern* des AA. Die Förderung von Sportbeziehungen (bspw. mit osteuropäischen Staaten, China, der Mongolei), die ebenfalls zur Sportförderung i.R. der auswärtigen Kulturpolitik des AA gehören, sind nicht mit einbezogen.

Aufgrund der zunehmenden Kritik an der sog. „Entwicklungshilfe“ in den 1990er Jahren²¹ suchte die Politik nach neuen Lösungen und Ansatzpunkten und Sport wurde von anderen Staaten als Instrument entdeckt.²² Neue Partner*innen²³ und die Popularität von Sport in vielen Partnerländern bewegten erste Regierungen dazu, Sport in ihr Portfolio aufzunehmen.²⁴ Die norwegische (seit 1983), kanadische (seit 1991), und die holländische Regierung (seit 1998) werden als Vorreiter bezeichnet, was die Nutzung von Sport in der EZ anbelangt (van Eekeren et al., 2013, S. 19ff.). Das *Sport for all*-Programm in Tansania der norwegischen EZ-Agentur in Kooperation mit dem NOK gilt als erster Meilenstein (vgl. ebd., S. 22). Die deutsche Bundesregierung setzte sogar bereits in den 1970er Jahren in unterschiedlichen Ländern Programme mit Schwerpunkt auf den Breitensport oder *Sport für alle* um (vgl., Giebenhain, 1990, S. 85; Digel & Fornoff, 1989, S. 4). Laut Digel und Fornoff (1989, S. 6) zählt „die Bundesrepublik zu den Staaten mit dem breitesten Spektrum und dem größten finanziellen und materiellen Volumen an Sportfördermaßnahmen“. Dass Deutschland dennoch international nicht als Vorreiter wahrgenommen wird, mag an der ‚Zurückhaltung‘ Anfang der 2000er liegen. Norwegen, die Schweiz, Kanada und Australien waren zu diesem Zeitpunkt jene Länder, die Sport zunehmend zur Erreichung ihrer entwicklungspolitischen Zielsetzungen nutzten (vgl. Levermore & Beacom, 2009, S. 3).

Ende der 1980er Jahre wurde dem Sport in der dt. EZ noch *keine Legitimation* zugesprochen (Digel & Fornoff, 1989, S. 353). Dennoch wurde empfohlen, „[...] dass die Entwicklungszusammenarbeit auf dem Gebiet des Sports auch für die Zukunft als entwicklungs-

²¹ S. hierzu auch Kapitel 4.3 dieser Arbeit.

²² Laut van Eekeren et al. (2013) war auch eine zeitgleiche nationale Bewegung von Bedeutung, die u.a. in den Niederlanden und Dänemark stattfand, nämlich die Wiederentdeckung und Anerkennung der *sozialen Rolle* des Sports.

²³ Durch die Veränderungen in der EZ-Landschaft waren ab dato auch neue Akteur*innen involviert: Neben Regierungen, multilateralen Institutionen und traditionellen Entwicklungsorganisationen werden zivilgesellschaftliche Akteur*innen/NGOs zu einer neuen, vierten Säule in der EZ (Coalter, 2009; Darnell & Black, 2011; van Eekeren et al., 2013). Hinsichtlich des Sports bedeutet dies, dass Organisationen aus dem nichtstaatlichen Bereich, die Sport i.R. ihrer Arbeit nutzen, in den 1990ern einen großen Zuwachs erfahren (bspw. Right to Play/Kanada, Magic Bus/Indien; Emima/Kenia) und auch Sportorganisationen zu neuen, attraktiven EZ-Partner*innen wurden (Coalter, 2009).

²⁴ Zeitgleich verankerten UN-Agenturen, internationale Sportverbände sowie internationale und nationale NGOs zunehmend Sport in ihren Policies und Agenden (van Eekeren et al., 2013; Levermore & Beacom, 2009).

politischer Bereich erhalten und [...] ausgebaut werden sollte“ (ebd., S. 9). Trotz rückläufiger Finanzmittel (vgl. Tab. 1) wurden ab den 1990ern einige Maßnahmen durchgeführt, bei denen allerdings hauptsächlich die *Entwicklung von Sportarten* im Vordergrund stand: „Nach dem fast vollständigen Ausstieg des BMZ bis Mitte der 90er Jahre [...] werden nun alle Projekte der GTZ aus dem Fond des AA finanziert. Hinsichtlich des Antragsverfahrens, der Konzeption, der Dauer und der Inhalte unterscheidet sich die Durchführung der Vorhaben nur noch geringfügig von denen des NOK [...]“ (Hinz, 2001). In einer GTZ-Auswertung von Eigenmaßnahmen der Kinder- und Jugendförderung im Zeitraum von 1994-2001 wurde der Einsatz von Sport unter „innovativen Ansätzen und Methoden“ empfohlen (Löwen, 2001, S. 28).²⁵ Dennoch dauerte es noch über zehn Jahre, bis dem Sport in der dt. EZ eine größere Rolle zugesprochen wurde.

Mit der Formulierung der *Millenniums-Entwicklungsziele (MDGs)* 2000 und 2001 veröffentlichten die Vereinten Nationen gleichzeitig die Programmlinie *Sport, Entwicklung und Frieden (Sport for Development and Peace – SDP)*. Von da an reihten sich die Meilensteine bzgl. der Etablierung des Sports in der internationalen Zusammenarbeit aneinander.²⁶ Während Regierungen anderer Länder den *Call for Action der Magglingen Deklaration* von 2003 wörtlich nahmen und Sport zu einem „wesentliche[n] Bestandteil der nationalen und internationale Entwicklungszusammenarbeit“ machten (DEZA, 2005, S. 11), waren die Finanzmittel für die Sportförderung des BMZ rückläufig, bevor sie ab 2005 komplett eingestellt wurden (s. Tab. 2). Im 11. Sportbericht der Bundesregierung wird erläutert:

„Angesichts der Notlage vieler Entwicklungsländer, insbesondere der am wenigsten entwickelten, und der aufgrund der vorhandenen Haushaltsmittel notwendigen Schwerpunktsetzung orientiert sich die deutsche Entwicklungszusammenarbeit vorrangig an der Unterstützung von Maßnahmen zur Armutsbekämpfung, Bildung, gesundheitlichen Basisversorgung und Bekämpfung von HIV/Aids sowie des Umweltschutzes. Sportförderung spielte in diesem Zusammenhang nur eine untergeordnete Rolle und findet seit 2005 nicht mehr statt.“ (Dt. Bundestag, 2006, S. 101)

²⁵ Drei der analysierten Maßnahmen hatten explizite Sportschwerpunkte: Ein überregionales Programm zur Rehabilitation von behinderten und kriegsversehrten Kindern und Jugendlichen, ein Projekt in Indien um Jugendliche durch Spiel- und Sportaktivitäten zu mobilisieren und ein Projekt zur Förderung des Breiten-, Freizeit- und Gesundheitssports in Jordanien (vgl. Löwen, S. 31f.; 55f.).

²⁶ Auf eine ausführliche und vollständige Aufführung der Meilensteine im Bereich SfE ab 2000 wird an dieser Stelle verzichtet. S. hierzu bspw. Brand (2020) oder van Eekren et al. (2013, S. 22ff.).

Dennoch werden hier sportbezogene BMZ-Projekte im Bericht aufgeführt, für die zwischen 2003 und 2006 rund 500 000 Euro zur Verfügung gestellt wurden (ebd.; s. Tab. 2). Auch wenn im Berichtszeitraum des 12. Sportberichts (2006-2009) keine sportbezogenen Gelder im BMZ-Haushalt eingeplant waren, fanden v.a. i.R. der Weltmeisterschaft (WM) 2006 in Deutschland sowie in Vorbereitung auf die WM in Südafrika einige Maßnahmen statt, für die das BMZ bis 2012 insgesamt rund 20 Mio. Euro investierte (Dt. Bundestag, 2010, s. Tab. 2). Neben dem Bau von Bolzplätzen wurde ab 2007 das bis dahin größte sportbezogene Entwicklungsprojekt *Youth Development through Football (YDF)* mit 7,5 Mio. Euro (bis 2012) gefördert. Mit einer Gesamtlaufzeit von sieben Jahren (04/2007 bis 03/2014) implementierte die GTZ/GIZ das Vorhaben in Südafrika und in neun weiteren afrikanischen Ländern (GIZ/YDF, 2010, 29. Juni). Zielgruppe waren benachteiligte junge Menschen aus städtischen und ländlichen Armutsgebieten, die über Sportprogramme gewalt- und gesundheitspräventive (bspw. zu Risiken von Alkohol und Drogen sowie von HIV/Aids) sowie geschlechtersensible Verhaltensweisen erlernen sollten. Innerhalb der Projektlaufzeit wurden zahlreiche Produkte (bspw. YDF Manuale zu unterschiedlichen Themenschwerpunkten) und Aktivitäten (bspw. Aus- und Weiterbildungsmodelle) entwickelt und implementiert. Sie wurden im weiteren Verlauf als Grundlage für nationale und international sportbezogene Entwicklungsprojekte genutzt (Petry, Bauer, Dumon & Wyszynski, 2015).

Insgesamt gab es in den Nullerjahren nur wenige sportbezogene EZ-Projekte. Die Schwerpunkte lagen u.a. auf der Friedenserziehung und Konfliktprävention,²⁷ auf der Inklusion von Jugendlichen mit Behinderungen²⁸ sowie auf Bildungsaspekten (bspw. Gesundheit, Gender) i.R. von Sportgroßveranstaltungen²⁹. Einige der Projekte wurden i.R. des ersten Auftrags mit der DSHS dokumentiert und hinsichtlich ihrer gemachten Erfahrungen ausgewertet, wodurch Herausforderungen identifiziert und Handlungsempfehlungen für die künftige Praxis gegeben werden konnten (vgl. Petry et al., 2015).

²⁷ Bspw. Philippinen (Laufzeit 2005-2010), Kongo (2005-2011), Kolumbien (2007-2015); Timor-Leste (2008-2011), Sri Lanka (2011) (vgl. Petry et al., 2015).

²⁸ Bspw. überregionales Projekt zur Rehabilitation kriegsversehrter Jugendlicher durch sportliche Aktivitäten in Angola, Kambodscha und El Salvador (2001-2014) (vgl. Petry et al., 2015)

²⁹ Z.B. zu HIV/Aids-Prävention i.R. der EM 2012 in der Ukraine (Laufzeit 2010-2012), Überregionales Jugendprojekt zu HIV/Aidsprävention, Gender u.a. Themen i.R. der WM in Südafrika (2007-2014) (vgl. Petry et al., 2015).

Die folgende Tabelle fasst die Fördermittel für die internationale Sportförderung des AA und BMZ zusammen. Während die Fördermittel des AA sich zwischen rund 3-5 Mio. Euro bewegten, standen dem BMZ über fast zwei Wahlperioden (2005-2011) keine sportbezogenen Fördermittel zur Verfügung. Ab 2013 stiegen die Fördermittel des BMZ sprunghaft an: Die GIZ-SfE-Vorhaben der GIZ wurden insgesamt mit rund 50 Mio. Euro ausgestattet (vgl. Dt. Bundestag, 2019, S. 157).³⁰

Tab. 2: Fördermittel der Bundesregierung für die Internationale Sportförderung des AA und BMZ 2000-2017 (Dt. Bundestag, 2006, 2010, 2014, 2019)

Jahr	AA ³¹	BMZ	Anderweitige (keine direkte Sportförderung)	BMZ-Finanzierungen
2000	4 300 000 DM	656 000 DM		
2001	4 800 000 DM	392 000 DM		
2002	2 647 000 €	138 000 €		
2003	2 647 000 €	91 000€	- 101 566 € Entwicklung einer globalen Kommunikations- und Kooperationsplattform im Internet im Bereich Fußball und Entwicklung (2003)	
2004	2 959 000 €	40 000€		
2005	2 775 000 €	0 €		
2006	2 768 000 €	0 €	- 275 000 € „WM Schulen 2006, Fair Play for Fair Life (2004-2006) - 100 000 € Jugendaustauschprogramm „Football for Development“ (2005)	
2007	2 825 000 €	0 €	- 140 000 € Schulwettbewerb „Lebenswelten Jugendlicher in der Einen Welt“ (05/2009-12/2010)	
2008	4 517 000 €	0 €		
2009	5 272 000 €	0 €	- 517 430 € Projekt „Fußball, Lernen, Global“ (04/2007-12/2009)	
2010	4 679 000 €	0 €		
2011	4 230 000 €	0 €	- 6 600 000 € Host-City-Programm Südafrika (2007-2010)	
2012	3 975 000 €	100 000 €	- 7 500 000 € YDF-Vorhaben (2007-2012) - 4 500 000 € Bau von Bolzplätzen Programm „Gewaltprävention durch Kinder- und Jugendfußball“ (2010-2012)	
2013	3 828 000 €	1 550 000 €		
2014	3 100 000 €	9 750 000 €		
2015	3 800 000 €	10 000 000 €		
2016	4 400 000 €	21 500 000 €		
2017	4 100 000 €	3 000 €		

Folgende Abbildung fasst ebenfalls die finanziellen Zuwendungen seit Förderbeginn zusammen und veranschaulicht dabei die *starken Schwankungen* im Hinblick auf die sportbezogene EZ. Auch wenn der 14. Sportbericht der Bundesregierung (Dt. Bundestag, 2019) nur die Zuschüsse bis 2017 festhält, wird angenommen, dass der BMZ-Trend auch

³⁰ Das SV SfE mit 27 Mio. Euro (2013 bis Ende 2018), das RV S4DA mit 16 Mio. Euro (2014-2018) und das RV Jordanien und Irak mit 4 Mio. Euro (2016-2019) (Dt. Bundestag, 2017, S. 157).

³¹ Die Zahl bezieht sich lediglich auf die *Sportförderung in den Entwicklungsländern* des AA. Die Förderung von Sportbeziehungen (bspw. mit osteuropäischen Staaten, China, der Mongolei), die ebenfalls zur Sportförderung i.R. der auswärtigen Kulturpolitik des AA gehören, sind nicht mit einbezogen.

in den Folgejahren bis heute (Stand: Mai 2022) weiter gestiegen oder auf einem ähnlich hohen Niveau geblieben ist.

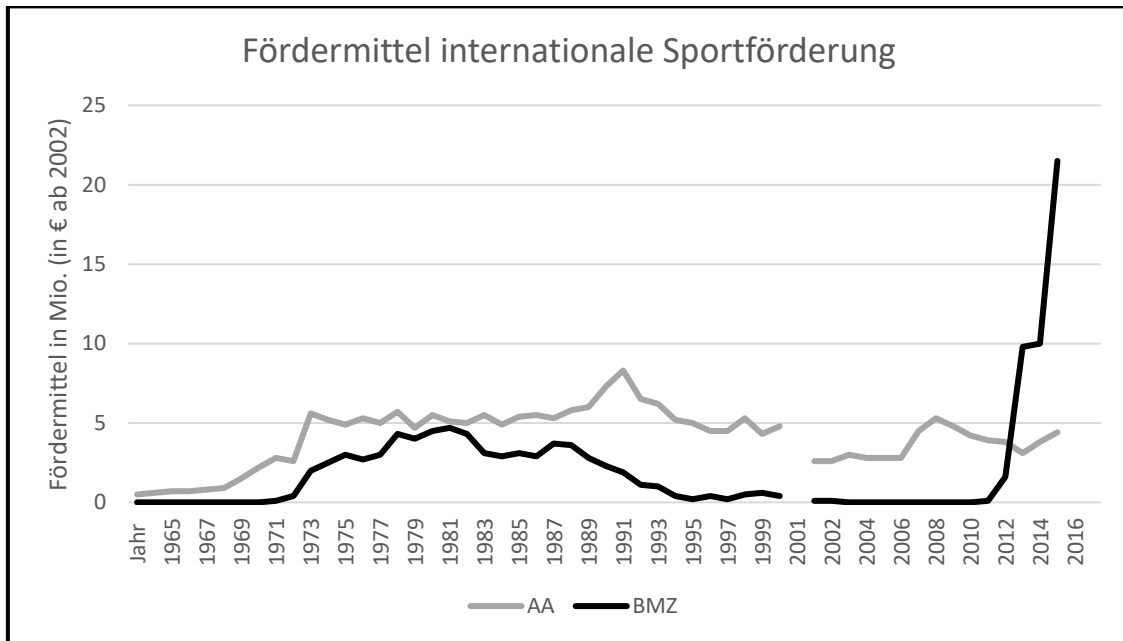


Abb. 1: Fördermittel internationale Sportförderung AA und BMZ im Zeitverlauf (eigene Darstellung)

2.2 Neupositionierung von Sport in der dt. EZ: Sport für Entwicklung ab 2013

Mit der Einrichtung des SVs SfE im Jahr 2013 misst das BMZ dem sportbezogenen EZ-Ansatz besondere Aufmerksamkeit bei und „Sport als Instrument zur Erreichung von Entwicklungszielen [ist] ein programmatischer Teil des entwicklungspolitischen Portfolios der Bundesregierung“ (Klingmann, Mialkas & Weinberg, 2020, S. 144). SfE ist in der dt. EZ zu einem übergreifenden Thema mit zahlreichen Verbindungen zu verschiedenen entwicklungspolitischen Zielen der Bundesregierung geworden (BMZ, 2010-2020a). Durch die Beauftragung der GIZ mit der Durchführung verschiedener Vorhaben (SV SfE seit 2013, seit 2014 Regionalprogramm „1000 Chancen für Afrika“, seit 2016 Regionalprogramm „Jordanien und Nordirak“), lässt sich ein neuer Schwerpunkt in der Umsetzung von SfE in der dt. EZ erkennen (Schmidt & Arnemann, 2020).

2.2.1 Vorhaben und Projekte der GIZ

Sektorvorhaben ‚Sport für Entwicklung‘ (SV SfE)

Die übergeordnete Zielsetzung des SVs SfE in der *ersten Phase (2013-2016)* war die verstärkte Nutzung von SfE als sektorübergreifendes Instrument in der EZ zur Erreichung entwicklungspolitischer Ziele. Das *Ziel der verstärkten Nutzung und Verbreitung von SfE* bleibt auch über die kommenden Phasen hinweg bis Ende 2022 erhalten.

Im Rahmen der ersten Phase wurden Pilotprojekte in sieben Ländern (Afghanistan, Brasilien, Jordanien, Kolumbien, Mosambik, Namibia und den Palästinensischen Gebieten) mit unterschiedlichen Schwerpunkten initiiert.³² Hierbei sollten unterschiedliche (strategische) SfE-Ansätze und Methoden³³ gemeinsam mit ausgewählten Partnerorganisationen entwickelt und getestet werden. Bis Ende der Laufzeit wurden in den Ländern

³² Zudem wurde zeitweise auf Initiative von CARE eine Maßnahme in Bosnien unterstützt (vgl. S4D Resource Toolkit, o.D.a unter Bosnia).

³³ Im Rahmen von (strategischen) *Ansätzen* geht es darum, welche überordneten Prozesse (Frage nach „was?“) in jedem Pilotland durchgeführt werden, um damit einen Beitrag zur Erreichung der Entwicklungsziele (SDGs) zu verfolgen (Bsp. Strategischer Ansatz in den Palästinensische Gebiete: Sport als Medium zur Berufsbefähigung). Mit *Methoden* sind konkrete Vorgehensweise/die Umsetzung eines Ansatzes (Frage nach „wie“: Wie wird Sport zur Erreichung der Entwicklungsziele eingesetzt?) gemeint. Bsp.: In Kolumbien wurde die Methode „Fútbol con principios“ basierend auf dem länderspezifischen Ansatz entwickelt. Die Methode ist Gegenstand des Manuals und wird i.R. der Ausbildungsworkshops an Trainer*innen weitergegeben (Petry & Bauer, 2016, S.II)

rund 75 SfE-Weiterbildungsworkshops für ca. 1000 Trainer*innen, Lehrkräfte und Sozialpädagog*innen durchgeführt. Die Etablierung des *Thementeam*s SfE in Deutschland während dieser Phase ist ein wichtiger Meilenstein bzgl. der Positionierung des Themas bei nationalen Akteur*innen sowie deren Vernetzung und verstärkter Zusammenarbeit. Die Sitzungen finden regelmäßig ein- bis zweimal jährlich statt und garantieren den Austausch zwischen staatlichen Akteur*innen (AA, BMI, BMZ), Durchführungsorganisationen (GIZ, Engagement Global), Verbänden (DOSB, DFB), *Non-Governmental Organizations* – NGOs (Right to Play, Streetfootballworld) und Akteur*innen aus der Wissenschaft (DSHS, International Council of Sport Science and Physical Education - ICSSPE) (Petry & Bauer, 2016).³⁴ „Nichtstaatliche Akteure, [...] nationale und regionale Sportverbände [...], sowie Vereine und Stiftungen gestalten somit den Ansatz SfE in Deutschland, aber auch in Partnerländern der deutschen EZ mit“ (Klingmann et al., 2020, S. 146f.).

In der *zweiten Phase (2016-2019)* galt es, die bereits entwickelten und erprobten SfE-Ansätze und -Methoden thematisch ausdifferenzieren (bspw. hinsichtlich ihrer Beiträge zu den SDGs), zu verfeinern und im Sinne des *Mainstreamings* zu verbreiten. Neue Partnerschaftsmodelle (bspw. mit Vereinen, Verbänden, NGOs, mit der Wirtschaft) wurden ausgebaut und neu initiiert. Die Arbeit in den Partnerländern richtete den Fokus auf die *Scaling-up-Prozesse* bzgl. der SfE-Ansätze, d.h. in Ländern aus der ersten Phase sollten die Ansätze in die Partnerstrukturen überführt werden bzw. darin verankert sein, in neuen Partnerländern (bspw. Indonesien, Region West-Balkan, Uganda)³⁵ galt es, diese Prozesse ‚einzuleiten‘.³⁶ In fünf Ländern sind die SfE-Ansätze seit Ende der Phase II fester Bestandteil der Partnerstrukturen (bspw. in Brasilien bei sieben NGOs eines nationalen Netzwerks). Außerdem waren Ende der Phase durch 300 Workshops rund 3400 Coaches und 280 Instructors geschult.³⁷ Ein weiterer wichtiger Schritt für das Mainstreaming von

³⁴ Die Akteur*innen werden im folgenden Kapitel genauer vorgestellt.

³⁵ Zudem wurde eine SfE-Komponente i.R. des GIZ-Bildungsprogramms für syrische Flüchtlinge und aufnehmende Gemeinden“ in der Türkei geschaffen (vgl. Tab. 3 und Birdsall, 2019).

³⁶ Zu den *Scaling-up-Prozessen* wurde 2017 in Zusammenarbeit mit der DSHS ein länderübergreifendes Standardindikatorenset auf Basis der GIZ Scaling-up Schlüsselfaktoren und dazugehörige Erhebungsinstrumente entwickelt. S. hierzu Ausführungen in Kap. 6.3.

³⁷ In der zweiten Phase wurde eine verstärkte Unterscheidung zwischen verschiedenen Zielgruppen in den SfE-Fortbildungen vorgenommen. In manchen Ländern ist die Unterscheidung zwischen Coach und Instructor allerdings nicht ganz eindeutig, da Personen sowohl mit Kindern und Jugendlichen arbeiten, als auch in der Erwachsenenbildung tätig sind (Petry & Bauer, 2019).

SfE ist die Entwicklung der *Homepage S4D Resource Toolkit*³⁸, die Ende 2018 der Öffentlichkeit präsentiert wurde und online ging. Hierfür wurden vorhandenes Wissen bzw. Lernerfahrungen aus der Praxis der Pilotmaßnahmen gebündelt und aufbereitet. Das Toolkit enthält ca. 300 einführende, themen- oder länderspezifische Dokumente (u.a. Factsheets, Publikationen, *Monitoring und Evaluations – M&E*-Instrumente, Leitfäden, Manuale etc.), die einer breiten potenziellen Nutzergruppe (Ministerien, EZ-Mitarbeiter*innen, Praktiker*innen) zur Verfügung stehen (Petry & Bauer, 2019).

In der *dritten Phase des SV SfE (2019-2022)* wird hinsichtlich der übergeordneten Zielsetzung der Fokus auf die verstärkte Anwendung von SfE-Ansätzen bei anderen EZ-Vorhaben und Stakeholdern aus der Privatwirtschaft und aus dem Sport (bspw. Bundesligaklubs, Verbände etc.) sowie auf die Messung von Wirkungen gesetzt. D.h. es geht nach zahlreichen Jahren der Implementierung um den *Nachweis*, inwiefern Sport entwicklungspolitische Ziele unterstützen und welche Wirkungen bei Kindern und Jugendlichen durch SfE-Maßnahmen ausgelöst werden können. Zudem werden die SfE-Maßnahmen auf andere Reformpartnerländer des BMZ ausgeweitet, bspw. in Marokko und Tunesien.³⁹ Alle Aktivitäten der dritten Phase sind seit Beginn 2021 stark durch die *Corona-Pandemie* betroffen: Aufgrund der weltweiten Bekämpfungsmaßnahmen (bspw. Lockdowns, Schulschließungen, Hygienemaßnahmen etc.) können geplante Fortbildungen nicht stattfinden und Kinder und Jugendliche werden nicht erreicht. Zahlreiche überbrückende und kreative Initiativen der Partnerländer, bspw. Online-Fortbildungen, Videos mit kleinen Aktivitäten oder die Entwicklung eines SfE-Hygienemanuals in Uganda, ersetzen die Arbeit mit Gruppen ‚auf dem Platz‘ allerdings nicht.

³⁸ S. S4D Resource Toolkit (o.D.).

³⁹ Partnerschaften mit sog. *Reformländern* „[...] fordern Eigenleistung auf der Basis des Ansatzes des Förderns und Forderns. Ein wichtiger Bestandteil ist dabei der Aufbau eigener Finanz- und Steuerverwaltungen, von Rechnungshöfen und Institutionen zur Bekämpfung von Korruption“ (BMZ, 2018, S. 6).

Regionalvorhaben ‚Sport für Entwicklung in Afrika‘ (RV S4DA)

Durch die Gründung der *Sonderinitiative „Mehr Platz für Sport – 1000 Chancen für Afrika“* durch Bundesminister Gerd Müller wurde Ende 2014 ein weiteres Vorhaben, das RV S4DA, in der GIZ etabliert. Neben deutschen Verbänden (DOSB, DFB) ist auch die FIFA Partner*in der Initiative, i.R. derer u.a. Sportplätze in ausgewählten afrikanischen Ländern rehabilitiert oder neugebaut werden (GIZ RV S4DA, Mai 2017; GIZ SV SfE, November 2016). Vor diesem Hintergrund entwickelte das RV S4DA in Zusammenarbeit mit lokalen und internationalen Partner*innen „Sportplätze mit Konzept!“. D.h., dass zum einen verantwortliche Institutionen (z.B. Schulen, Kommunen) der neu gebauten oder renovierten Sportplätze zu deren nachhaltiger Nutzung und Pflege beraten und geschult werden. Zum anderen werden Multiplikator*innen zu SfE-Ansätzen und Methoden⁴⁰ fortgebildet, die für Kinder und Jugendliche auf den Plätzen SfE-Aktivitäten anbieten. „Durch die Kombination von Organisationsberatung, der Bereitstellung von Infrastruktur und der Qualifikation von Einzelpersonen entsteht ein Gesamtkonzept zur nachhaltigen Umsetzung von SfE“ (Schmid & Arnemann, 2020, S. 163). Das Vorhaben ist in 13 Ländern Subsahara-Afrikas präsent. Mit dem Ziel, kontextunabhängige Standards zur Umsetzung von SfE zu designen, wurde bewusst ein *Mehr-Länder-Vorhaben* geplant (ebd.).⁴¹ Mit der Unterstützung der DSHS etablierten sich frühzeitig länderübergreifende Monitoring-Standards, die zum *Wirkungsorientierten Monitoring (WoM)* beitragen und aussagekräftige, aggregierbare Daten erzeugen, die für die Auftraggebenden wichtig sind: Von den Betreiber*innen der renovierten/neu gebauten Sportplätze in 13 Ländern verfügen fast 90 Prozent über Konzepte, in denen SfE verankert ist. Fast alle der 620 zusätzlich geschulten SfE-Multiplikator*innen bieten ein pädagogisch-wertvolles SfE-Training nach definierten Standards an. Von den Angeboten auf den Plätzen profitieren mehr als 490000 Kinder und Jugendliche, rund 9400 davon nehmen regelmäßig an angebotenen SfE-Trainings teil (vgl. ebd., S. 161f.).

⁴⁰ Je nach Bedarf der lokalen Partner*innen wird „Sport dabei zur Förderung von Bildung, beruflicher Bildung, friedlichem Zusammenleben, Gesundheit (HIV/AIDS Prävention) oder zur Förderung der sozialen und politischen Teilhabe eingesetzt“ (Schmid & Arnemann, 2020, S. 155).

⁴¹ Eine Sammlung der bereits entwickelten Standards (bspw. Orientierungshilfen für Partner*innen zur Integration von Sport in bereits bestehende Konzepte, Handreichungen zur Instandhaltung von Sportplätzen, sportpädagogische Handbücher für Multiplikator*innen etc.; vgl. Schmid & Armenann, 2020, S. 161) ist im S4D Resource Toolkit (o.D.) zu finden.

Regionalvorhaben (RV) Jordanien und Irak

Seit der Flüchtlingskrise 2015 sind die *Minderung von Fluchtursachen*, die *Stabilisierung von Aufnahmeregionen* und die *Unterstützung von Flüchtlingen* (Integration und Reintegration) zentrale Themen der deutschen Bundesregierung. Allein 2016 wurden hierfür drei Milliarden Euro zur Verfügung gestellt (BMZ, 2010-2020b). Im Rahmen dieses Engagements wird unter dem Titel „Austausch, Bildung und Konfliktbearbeitung durch Sport für Entwicklung“ durch die GIZ seit 2016 (bis voraussichtlich 2024) ein weiteres RV in Jordanien und Irak umgesetzt. Hier werden gemeinsam mit lokalen Partner*innen inklusive sportliche Aktivitäten zur „körperlichen und seelischen Gesundheit“ und mit dem Ziel der Konfliktvermeidung in irakischen und jordanischen Flüchtlingslagern sowie in Aufnahmegemeinden angeboten. In vier jordanischen Regionen, in denen besonders viele Flüchtlinge leben, sowie in sechs Flüchtlingslagern und zehn Jugendzentren vier irakischer Distrikte gibt es SfE-Aktivitäten⁴² in über 150 Schulen und 100 Sportstätten (Nölting, 2018). Im Rahmen dieses RVs wurden zahlreiche pädagogische Handbücher zu unterschiedlichen Sportarten entwickelt,⁴³ allerdings ist insgesamt nur wenig öffentlich dokumentiert.

Alle Vorhaben arbeiten in den Partnerländern i.R. ihrer *Capacity Development (CD)-Maßnahmen*⁴⁴ im Sinne eines *Mehr-Ebenen-Ansatzes*: Auf Ebene der *Gesellschaft* werden bspw. nationale und regionale Ministerien (oft im Bereich Bildung, Kultur oder Sport), Bildungseinrichtungen (bspw. Universitäten), Gemeinden und Sportverbände hinsichtlich der nachhaltigen Verankerung von SfE in ihren nationalen Programmen, Lehrplänen, Lizenzen und Strategien beraten. Auf Ebene der *Organisation* geht es um das Thema *Organisationsentwicklung*. Die Vorhaben beraten hierbei „Partner aus Bildung, Sport und Zivilgesellschaft (z.B. Berufsschulen, Schulen, Jugend- oder Sportorganisationen) zur Verbesserung und Erweiterung ihrer Angebote durch die nachhaltige Integration von Sport für Entwicklung“ (Schmid & Arnemann, 2020, S. 159). Auf der Ebene

⁴² Darunter auch das Busprojekt „Yalla Riyada Bus“, das über Sportfestivals und Informationsangebote das Thema SfE in Kommunen bekannt machen soll (Wirsich, 2018).

⁴³ S. S4D Resource Toolkit (o.D.).

⁴⁴ „Capacity Development beschreibt einen Prozess, durch den Menschen, Organisationen und Gesellschaften ihre Fähigkeiten mobilisieren, anpassen und ausbauen, um ihre eigene Entwicklung nachhaltig zu gestalten und sich an verändernde Rahmenbedingungen anzupassen“ (GIZ, 2013, S. 3).

der *Person* werden über Fortbildungskurse Kompetenzen von Trainer*innen, Lehrkräften und Sozialarbeiter*innen gestärkt, sodass diese als SfE-Multiplikator*innen qualitativ hochwertige SfE-Aktivitäten für benachteiligte Kinder und Jugendliche anbieten können (ebd.). Zudem arbeiten alle Vorhaben mit *Multiakteursansätzen*, d.h. es werden intersektoral Partnerschaften mit Akteur*innen aus unterschiedlichen Bereichen (Politik, Wirtschaft, organisierter Sport, Zivilgesellschaft, Wissenschaft) geknüpft (vgl. Klingmann et al., 2020, S. 146)

Tabelle 3 gibt einen Überblick zu allen Ländern, in denen die deutsche EZ seit Gründung des SVs 2013 i.R. ihrer SfE-Maßnahmen tätig ist. Sie differenziert zwischen den zuständigen GIZ-Vorhaben und gibt weitere Informationen zu Laufzeiten, den spezifischen SfE-Ansätzen inklusive des Beitrages zu den *Nachhaltigen Entwicklungszielen (Sustainable Development Goals – SDGs)*, Zielgruppen, Kooperationsstrukturen, CD-Maßnahmen, Baumaßnahmen und weiteren inhaltlichen Schwerpunkten (bspw. Verankerung von Ansätzen in Bildungssystemen, Details zu Projektkontexten etc.) in den Ländern.

Tab. 3: Übersicht der Partnerländer der SfE-Vorhaben der GIZ (eigene Darstellung basierend auf: Birdsall, 2019; GIZ RV S4DA, Mai 2017a-e; GIZ SV SfE November 2017a-d; GIZ SV SfE, November 2022; Nölting, 2018; Petry & Bauer, 2016, 2019; S4D Resource Toolkit, o.D.a)

Land	Vorhaben & Laufzeit	SfE-Ansatz, Zielgruppe(n), Sportart(en), Kooperationen, CD-Maßnahmen, Baumaßnahmen und weitere Schwerpunkte in den Partnerländern ⁴⁵
AFRIKA⁴⁶		
Äthiopien	RV S4DA, (11/2014-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung beruflicher Bildung durch Sport (Beitrag zu SDG 4, 8) - <i>Zielgruppe(n)</i>: Jugendliche (Vermittlung arbeitsmarktrelevanter Schlüsselkompetenzen, bspw. Teamfähigkeit) - <i>Sportart(en)</i>: Fußball, Basketball, Handball, Volleyball - <i>Kooperationen</i>: Äthiopische Berufsbildungsbehörde (Federal TVET Agency), Ministerium für Jugend und Sport, staatliche Berufsschulen, Don Bosco Mondo e.V., äthiopische Sportverbände für Fußball, Basketball, Volleyball, Handball - <i>CD-Maßnahmen</i>: Fortbildung von 448 Coaches und 72 Instructors in der Methode „Sport2Work“

⁴⁵ Nicht zu jedem Land liegen umfangreiche Informationen vor. Es wurden angegebene Quellen genutzt, es besteht jedoch keine Gewährleistung auf Richtigkeit und Vollständigkeit, wie bspw. bei den Aufzählungen zu Kooperationen.

⁴⁶ Das RV S4DA hat in den folgenden Ländern weitere Sportplätze gebaut: Demokratische Republik Kongo (8), Mali (1), Niger (1), Nigeria (2), Ruanda (3), Tansania (2), Zentralafrikanische Republik (1); Da in den Ländern keine weiteren, inhaltlichen Maßnahmen durchgeführt wurden, werden sie an dieser Stelle nicht gelistet.

		<ul style="list-style-type: none"> - <i>Baumaßnahmen</i>: Renovierung/Neubau von 48 Sportplätzen an Berufsschulen und Jugendzentren⁴⁷
Elfenbeinküste	RV S4DA (07/2019-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung beruflicher Bildung durch Sport (Beitrag zu SDG 4, 8) - <i>Zielgruppe(n)</i>: Jugendliche (Vermittlung arbeitsmarktrelevanter Schlüsselkompetenzen, bspw. Teamfähigkeit) - <i>Sportart(en)</i>: Fußball, Basketball, Handball, Volleyball - <i>Kooperationen</i>: Berufsbildungsinstitutionen - <i>CD-Maßnahmen</i>: Fortbildung von 10 Coaches und 3 Instructors in der Methode „Se démarquer/Standing out“ - <i>Baumaßnahmen</i>: Renovierung/Neubau von 6 Sportplätzen
Ghana	RV S4DA (07/2019-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: nicht bekannt (n.b.) - <i>Zielgruppe(n)</i>: n.b. - <i>Sportart(en)</i>: Fußball - <i>Kooperationen</i>: Right to Play - <i>CD-Maßnahmen</i>: Fortbildung von 50 Coaches, 1 Instructor - <i>Baumaßnahmen</i>: Renovierung/Neubau von 2 Sportplätzen
Kenia	RV S4DA (11/2014-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Gewaltprävention durch Sport im Fluchtcontext (Beitrag zu SDG 16) - <i>Zielgruppe(n)</i>: Kinder und Jugendliche in Flüchtlingscamps und aufnehmenden Gemeinden - <i>Sportart(en)</i>: alle Disziplinen - <i>Kooperationen</i>: streetfootballworld, Trans-Nzoia Youth Sports Association (TYSA), Moving the Goalposts (MTG), Horn of Africa Development Initiative (HODI), Auma Obama Foundation – Sauti Kuu, Don Bosco Mondo e.V., Kenianischer Fußballverband, Lotus Kenya Action for Development Organisation, (LOKADO), Seeds of Peace Africa (SOPA); Kenya Academy of Sport (KAS); GIZ-Vorhaben “Unterstützung für Flüchtlinge, insbesondere Rückkehrwillige, und aufnehmende Gemeinden in Kenia“ (SIF); UNHCR; UNICEF - <i>CD-Maßnahmen</i>: Fortbildung von 253 Coaches und 81 Instructors in der Methode „Violence Prevention through Football“ - <i>Weitere Schwerpunkte</i>: Einwöchige Sportcamps für einheimische und geflüchtete Kinder - <i>Baumaßnahmen</i>: Renovierung/Neubau von 21 Sportplätzen
Marokko	SV SfE Phase III (10/2020-09/2022)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung beruflicher Bildung durch Sport (Beitrag zu SDG 4, 8) - <i>Zielgruppe(n)</i>: Jugendliche - <i>Sportart(en)</i>: Martial Art, Basketball, Fußball, (bald Parcours, Breakdance) - <i>Kooperationen</i>: Ministère délégué auprès du Ministre des Affaires Étrangères, de la Coopération Africaine et des Marocains Résidant à l'Étranger (MDCMRE), Institut Royal de Formation des Cadres de la Jeunesse et des Sports, TIBU, commune partenaires de RECOSA - <i>CD-Maßnahmen</i>: Fortbildung von 116 Coaches und 14 Instructors - <i>Baumaßnahmen</i>: keine
Mosambik	SV Phase I (06/ 2013-06/2016); RV S4DA (06/2016-06/2019)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Stärkung von Kommunen und Demokratieförderung (Gute Regierungsführung) durch Sport - <i>Zielgruppe(n)</i>: Über ausgebildete Personen marginalisierte Kinder, Jugendliche - <i>Sportart(en)</i>: Fußball, Ballsportarten - <i>Kooperationen</i>: Sportwissenschaftliche Fakultät der Pädagogischen Universität Maputo (UP), Gemeinderat der Stadt Matola, Don Bosco

⁴⁷ Die Zahl umfasst bereits gebaute/renovierte und geplante Sportplätze sowie jene, die gerade gebaut/renoviert werden.

		<p>Mosambik, Vereinigung zur Förderung der Entwicklung Jugendlicher (Khandlelo), Sportverein Albazine (ADA), Bayerischer Fußballverband</p> <ul style="list-style-type: none"> - <i>CD-Maßnahmen:</i> Fortbildung von 244 Coaches und 80 Instructors in den Methoden „Ballschule“ und „Treino Social“ - <i>Weitere Schwerpunkte:</i> Stärkung von Sportorganisationen (Vereine, NGOs, Hochschule) in ihrer Organisations- und Strukturentwicklung; Förderung der Zusammenarbeit zwischen Universitäten, Kommunen und lokalen Vereinen; Integration von SfE in den Lehrplan der Universität - <i>Baumaßnahmen:</i> Renovierung/Neubau von 10 Sportplätzen
Namibia	SV SfE Phase I-III (06/2013-06/2019); RV S4DA (06/2016-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz:</i> Förderung reproduktiver Gesundheit und Bildung durch Sport (Beitrag zu SDG 3, 5) - <i>Zielgruppe(n):</i> Jugendliche, insbesondere Mädchen und junge Frauen - <i>Sportart(en):</i> Fußball, Basketball, Volleyball, Netball, Physical Education (PE) - <i>Kooperationen:</i> Namibisches Ministerium für Jugend, Nationale Dienste, Sport und Kultur, Namibisches Ministerium für Bildung, Kunst und Kultur, namibische Sportverbände für Fußball (NFA), Basketball, Volleyball und Netball, DOSB, DFB, Fußball- und Leichtathletik-Verband Westfalen (FLVW), Futouris e.V., TUI Care Foundation, FIFA, TSG Hoffenheim, NGO SCORE, Basketball Artist School (BAS), UNICEF - <i>CD- Maßnahmen:</i> Fortbildungen für 630 Coaches und 114 Instructors in den „Methoden Football/Basketball/Netball/Volleyball4Life“ und „PE4Life“ - <i>Weitere Schwerpunkte:</i> „Football4Environment“-Camps in Kooperation mit dem TSG Hoffenheim für ca. 100 Personen (SV SfE) - <i>Baumaßnahmen:</i> Renovierung/Neubau von 46 Sportplätzen
Senegal	RV S4DA (07/2019-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz:</i> Förderung beruflicher Bildung durch Sport (Beitrag zu SDG 4, 8) - <i>Zielgruppe(n):</i> Jugendliche (Vermittlung arbeitsmarktrelevanter Schlüsselkompetenzen, bspw. Teamfähigkeit) - <i>Sportarten(en):</i> Fußball, Basketball, Handball, Volleyball - <i>Kooperationen:</i> National Sports Institute Dakar, Berufsbildungszentrum Thiès, SOS-Kinderdörfer, Don Bosco Mondo e.V. - <i>CD-Maßnahmen:</i> Fortbildung von ca. 90 Coaches, 26 Instructors in der Methode „Se démarquer/Standing out“ - <i>Weitere Schwerpunkte:</i> Verankerung des SfE-Ansatzes und -Methode im PE-Lehrplan des National Sports Institute in Dakar - <i>Baumaßnahmen:</i> Renovierung/Neubau von 19 Sportplätzen
Togo	RV S4DA (11/2014-06/2019)	<ul style="list-style-type: none"> - <i>SfE-Ansatz:</i> Stärkung der Zivilgesellschaft und Förderung der sozialen und politischen Teilhabe durch Sport - <i>Zielgruppe(n):</i> Benachteiligte Gruppen (v.a. Frauen und Mädchen, die in das soziale und politische Gemeindeleben eingebunden werden sollen) - <i>Sportarten(en):</i> Fußball, Basketball, Handball, Volleyball - <i>Kooperationen:</i> Togoisches Ministerium für Kommunikation, Kultur, Sport und Allgemeinbildung; Kommunalverwaltungen Lomé, Sokodé, Kara, Kpalimé; Don Bosco Mondo e.V.; SOS Kinderdörfer weltweit e.V./Hermann-Gmeiner-Fonds Deutschland e.V.; GIZ-Vorhaben „Gute Regierungsführung und Dezentralisierung“ - <i>CD-Maßnahmen:</i> Fortbildung von 139 Coaches, 14 Instructors in den Methoden „Le sport/football/handball/basketball/volleyball au service de civisme et de la bonne citoyenneté“ - <i>Baumaßnahmen:</i> Renovierung/Neubau von 17 Sportplätzen
Tunesien	SV SfE Phase III	<ul style="list-style-type: none"> - <i>SfE-Ansatz:</i> n.b. - <i>Zielgruppe(n):</i> Jugendliche

	(08/2019-09/2022)	<ul style="list-style-type: none"> - <i>Sportart(en)</i>: Ballsportarten, Martial Arts, Trekking, Carving, Speedball, u.a. - <i>Kooperationen</i>: drei lokale NGOs, Ministerium für Jugend und Sport - <i>CD-Maßnahmen</i>: Fortbildung von 80 Coaches - <i>Weitere Schwerpunkte</i>: Verankerung des SfE-Ansatzes in das Programm „Vision sectorielle de la jeunesse“ des Ministeriums für Jugend und Sport - <i>Baumaßnahmen</i>: Renovierung/Neubau eines Sportplatzes
Uganda	SV SfE Phase II (05/2018-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Partnerschaften über <i>Athletics for Development (A4D)</i> (Beiträge zu den SDGs 3, 4, 5, 8, 16 und 17) - <i>Zielgruppe(n)</i>: benachteiligte Kinder und Jugendliche, Gemeinden - <i>Sportart(en)</i>: Leichtathletik, kleine Spiele - <i>Kooperationen</i>: Ministerium für Bildung und Sport, Ugandisches Olympisches und Paralympisches Komitee, Training and Resources Network (24 ugandische oder in Uganda akkreditierte Organisationen, darunter Ministerien, Sportverbände, NGOs, Universitäten, internationale Kooperationspartner*innen, Lehrerverbände), World Athletics, DOSB, Viva con Agua - <i>CD-Maßnahmen</i>: Fortbildung von 368 Coaches und 24 Instructors in der Methode A4D - <i>Weitere Schwerpunkte</i>: Verankerung des SfE-Ansatzes und der A4D-Methode im Lehrplan der Makerere Universität - <i>Baumaßnahmen</i>: keine
ASIEN		
Afghanistan	SV SfE Phase I, II (02/2014-10/2017)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung von Mädchen und jungen Frauen durch verbesserte Sportangebote an Schulen (Beitrag zu SDG 4, 5) - <i>Zielgruppe(n)</i>: Lehrerinnen, Mädchen, junge Frauen - <i>Sportart(en)</i>: Ballsportarten, PE, kleine Spiele - <i>Kooperationen</i>: Afghanisches Bildungsministerium, Bildungsabteilung der afghanischen Provinzregierungen, GIZ-Programm zur Förderung der Grund- und Sekundarbildung in Afghanistan (BEPA) - <i>CD-Maßnahmen</i>: Fortbildung von 121 Sportlehrerinnen von 70 Mädchenschulen, 10 Instructors - <i>Weitere Schwerpunkte</i>: Entwicklung und Etablierung eines Lehrplans für PE innerhalb des Grundschullehramt-Studiengangs an sog. <i>Teacher Training Colleges</i> - <i>Baumaßnahmen</i>: keine
Indonesien	SV SfE Phase II, III (05/2018 – 12/2021)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Fußball als Instrument zur Gesundheitsförderung und Gewaltprävention für pädagogisch-wertvolle und qualitative Schulen (Beitrag zu SDGs 3, 4, 5) - <i>Zielgruppe(n)</i>: Kinder an Grund- und Mittelschulen; Sportlehrerinnen - <i>Sportart(en)</i>: Fußball, kleine und traditionelle Spiele (für Grundschule) - <i>Kooperationen</i>: Indonesisches Ministerium für Bildung und Kultur, nationale und internationale Verbände (Asian Football Confederation, DFB, FIFA, Indonesischer Fußballverband (PSSI), Universitäten) - <i>CD- Maßnahmen</i>: Fortbildung von 535 Coaches und 35 Instructors - <i>Weitere Schwerpunkte</i>: Verankerung von SfE im indonesischen Bildungs- und Sportsystem durch die Erweiterung bestehender Curricula (Fußball D-Lizenz, Lehrplan Grundschule); SfE E-Learning Modul für angehende Sportlehrkräfte an Universitäten - <i>Baumaßnahmen</i>: Renovierung/Neubau von zwei Sportplätzen
Irak	RV Jordanien/Irak (04/2018-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung des psychischen und physischen Wohlbefindens, sozialer Zusammenhalt, Gewaltprävention durch Fußballaktivitäten (Beitrag zu SDG 4, 16) - <i>Zielgruppe(n)</i>: Kinder und Jugendliche in Flüchtlingslagern, Jugendzentren

		<ul style="list-style-type: none"> - <i>Sportart(en)</i>: Fußball, Ultimate Frisbee, kleine Spiele - <i>Kooperationen</i>: Universität von Dohuk/College of Physical Education, locale NGOs, Friends of Waldorf Education, War Child UK - <i>CD-Maßnahmen</i>: Fortbildung von 326 Coaches und 5 Instructors - <i>Weitere Schwerpunkte</i>: SfE-Aktivitäten in Flüchtlingscamps und aufnehmenden Gemeinden (in schulischen und außerschulischen Systemen), institutionelle Vernetzung im Flüchtlingskontext - <i>Baumaßnahmen</i>: keine
Jordanien	SV SfE Phase I (04/2018-06/2016); RV Jordanien/Irak (07/2016-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung des psychischen und physischen Wohlbefindens, sozialer Zusammenhalt, Gewaltprävention durch Fußballaktivitäten (Beitrag zu SDG 4, 16) - <i>Zielgruppe(n)</i>: Kinder und Jugendliche an 150 Schulen und 100 Sportstätten - <i>Sportart(en)</i>: Fußball, Basketball, Handball, Ultimate Frisbee, kleine Spiele, - <i>Kooperationen</i>: Bildungsministerium, Asian Football Development Project, NGO Generations for Peace und Paces, Jordanisches Olympisches Komitee, Jordanischer Fußballverband, Universität von Jordanien, Al al-Beyt Universität - <i>CD-Maßnahmen</i>: Fortbildung von 1093 Coaches und 38 Instructors - <i>Weitere Schwerpunkte</i>: SfE-Aktivitäten in Flüchtlingscamps und aufnehmenden Gemeinden (in schulischen und außerschulischen Systemen), institutionelle Vernetzung im Flüchtlingskontext, Projekt „Yalla Riyada Bus“ - <i>Baumaßnahmen</i>: keine
Palästinensische Gebiete	SV SfE Phase I & II (03/2014-03/2019)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Förderung beruflicher Bildung und Arbeitsmarktbefähigung durch SfE (Beitrag zu SDG 4, 8) - <i>Zielgruppe(n)</i>: schwer erreichbare und sozial benachteiligte Jugendliche (aus Flüchtlingslagern, Schulabbrecher und arbeitslose Jugendliche) - <i>Sportart(en)</i>: Fußball, Ultimate Frisbee, kleine Spiele - <i>Kooperationen</i>: DFB, Palästinensisches Ministerium für Arbeit, Lutheran World Federation, NGO Palestine Sports for Life, GIZ-Vorhaben „Förderung von Berufsbildung und Arbeitsmarkt“ - <i>CD- Maßnahmen</i>: Fortbildung von 146 Coaches und 3 Instructors in den Methoden „Kicking Competencies“ und „Learn to play – play to learn“ - <i>Weitere Schwerpunkte</i>: Integration von SfE in Berufsorientierungstage und Berufsbildungsangebote für über 3000 Jugendliche - <i>Baumaßnahmen</i>: Renovierung/Neubau eines Sportplatzes
Sri Lanka	SV SfE Phase II (05/2017-02/2019)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Sport zur Förderung sozialer Kohäsion und Versöhnung (Beitrag zu SDG 10, 16) - <i>Zielgruppe(n)</i>: Sportlehrkräfte, Kinder und Jugendliche - <i>Sportart(en)</i>: Fußball, Netball, kleine Spiele - <i>Kooperationen</i>: Integrationskomponente in Bildungsvorhaben „Education for Social Cohesion Programme“ der GIZ, fünf nationale Bildungskollegs mit PE-Schwerpunkt - <i>CD-Maßnahmen</i>: Fortbildung von 160 Coaches und 10 Instructors im Konzept „Sport for Social Cohesion and Reconciliation“ - <i>Baumaßnahmen</i>: keine
Türkei	SV SfE Phase II & DOSB (06/2016-05/2019)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Sport zur Stärkung des friedlichen Zusammenlebens und den interkulturellen Austausch (Beitrag zu SDG 10, 16) - <i>Zielgruppe(n)</i>: syrische Flüchtlinge und die Bevölkerung der Aufnahmegemeinden - <i>Sportart(en)</i>: Fußball, Tischtennis, Basketball, Leichtathletik, Martial Arts

		<ul style="list-style-type: none"> - <i>Kooperationen</i>: GIZ „Bildungsprogramm für syrische Flüchtlinge und aufnehmende Gemeinden“, NGOs in der syrisch-türkischen Grenzregion - <i>CD-Maßnahmen</i>: Fortbildung von 360 Coaches - <i>Baumaßnahmen</i>: keine
LATEINAMERIKA		
Brasilien	SV SfE-Phase I, II (2014-06/2018)	<ul style="list-style-type: none"> - <i>SfE-Ansatz</i>: Fußball als Instrument zur Förderung von Gesundheit, Bildung, Geschlechtergleichstellung, Umwelt und Gewaltprävention (Beitrag zu SDG 3, 4, 5, 13, 16) - <i>Zielgruppe(n)</i>: benachteiligte Jugendliche - <i>Sportart(en)</i>: Fußball, divers - <i>Kooperationen</i>: u.a. Brasilianisches Sportministerium, Goethe-Institut NGOs Bola pra Frente und Promundo; streetfootballworld, DFB, fünf NGOs des brasilianischen <i>Sport for Social Change</i> Netzwerks (portugiesisch: Rede Esporte pela Mudança Social – REMS) - <i>CD-Maßnahmen</i>: Fortbildung von 1185 Coaches und 30 Instructors in den Methoden „Treino Social I & II“ - <i>Weitere Schwerpunkte</i>: Veranstaltungen rund um die WM 2014 und die Olympischen Spiele 2016 (Camps für Kinder und Jugendliche, Festivals, Podiumsdiskussionen) - <i>Baumaßnahmen</i>: keine
Ecuador	SV SfE Phase III (11/2019-09/2022)	<ul style="list-style-type: none"> - <i>SfE-Ansatz und Zielgruppe(n)</i>: Friedensentwicklung und Gewaltprävention durch Sport (Beitrag zu SDG 3, 4, 16) - <i>Sportart(en)</i>: Fußball, diverse - <i>Kooperationen</i>: Sportsekretariat Ecuador, Unidad Educativa Particular Adventista del Ecuador (CADE), Consejo Iberoamericano del Deporte (CID) - <i>CD-Maßnahmen</i>: Fortbildung von 46 Coaches und 22 Instructors - <i>Baumaßnahmen</i>: keine
Kolumbien	SV SfE Phase I-III (01/2015-09/2022)	<ul style="list-style-type: none"> - <i>SfE Ansatz</i>: Friedensentwicklung und Gewaltprävention durch Sport (Beitrag zu SDG 3, 4, 16) - <i>Zielgruppe(n)</i>: Kinder und Jugendliche in Konfliktregionen, Binnenflüchtlinge, ehemalige Kämpfende - <i>Sportart(en)</i>: Fußball, diverse - <i>Kooperationen</i>: kolumbianisches Außenministerium, Ministerio del Deporte (Mindeporte), NGOs, DFB, DOSB, verschiedene kol. Gemeinden/Bildungssekretariate; GIZ-Vorhaben „Unterstützung der Friedensentwicklung in Kolumbien“ (PRO-PAZ) - <i>CD-Maßnahmen</i>: Fortbildung von 1790 Coaches und 150 Instructors in den Methoden „Fútbol/Deporte con Principios“ - <i>Weitere Schwerpunkte</i>: Gründung eines kol. Uni-Netzwerks zur Wirkungsmessung im Bereich SfE (in Kooperation mit der DSHS) - <i>Baumaßnahmen</i>: Renovierung/Neubau von 7 Sportplätzen
Paraguay	SV SfE-Phase II, III (01/2019-09/2022)	<ul style="list-style-type: none"> - <i>SfE-Ansatz und Zielgruppe(n)</i>: Friedensentwicklung und Gewaltprävention durch Sport (Beitrag zu SDG 3, 4, 16) - <i>Sportart(en)</i>: Fußball, diverse - <i>Kooperationen</i>: Sportsekretariat Paraguay, Consejo Iberoamericano del Deporte (CID) - <i>CD-Maßnahmen</i>: Fortbildung von 308 Coaches und 13 Instructors - <i>Baumaßnahmen</i>: keine

SÜD-/OST-EUROPA		
Region West-Balkan (5 Länder): Albanien, Bosnien und Herzegowina, Kosovo, Nord Mazedonien, Serbien	SV SfE Phase II, III (07/2017-aktuell)	<ul style="list-style-type: none"> - <i>SfE-Ansatz:</i> Sport, Bewegung und Spiel als Instrumente für soziale Inklusionsprozesse und „Community Development“ (Beitrag zu SDG 10, 16) - <i>Zielgruppe(n):</i> benachteiligte Gruppen (bspw. Roma), Kommunen - <i>Sportart(en):</i> Divers - <i>Kooperationen:</i> lokale Netzwerke in den Ländern (Kommunen, Schulen, Kindergärten, Sportvereine) - <i>CD- Maßnahmen:</i> Fortbildung von insgesamt 893 Coaches und 70 Instructors (Albanien: 257 Coaches, 18 Instructors; Bosnien und Herzegowina: 119 Coaches, 14 Instructors; Kosovo: 294 Coaches und 14 Instructors; Nord Mazedonien: 171 Coaches, 17 Instructors; Serbien: 52 Coaches, 7 Instructors) - <i>Weitere Schwerpunkte:</i> In jedem Land fanden i.R. der lokalen Netzwerke unterschiedliche Aktivitäten statt (s. S4D Resource Toolkit, o.D.a) - <i>Baumaßnahmen:</i> keine

2.2.2 Kooperationsstrukturen und deutsche SfE-Akteur*innen

Die zuvor aufgeführten *Kooperationen in den Partnerländern* (s. Tabelle 3) sind zahlreich und sehr unterschiedlich. Dennoch wird in Abbildung 2 versucht, gängige Kooperationsstrukturen in den Partnerländern der GIZ-Vorhaben abzubilden:

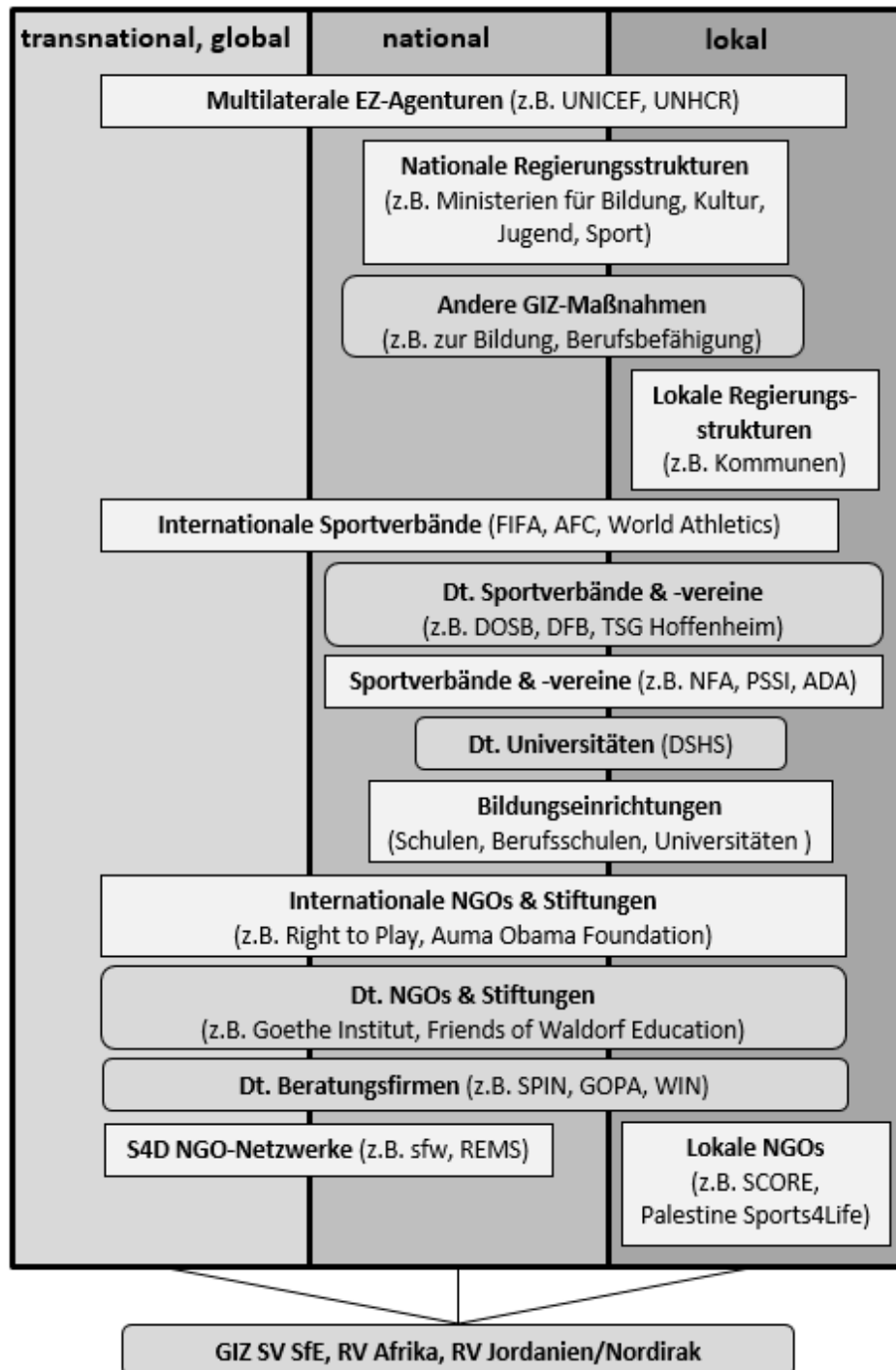


Abb. 2: Kooperationsstrukturen der GIZ-Vorhaben in den Partnerländern (eigene Darstellung in Anlehnung an Naish, 2017)

Das Schaubild unterscheidet zwischen Partner*innen und Kooperationen auf *lokaler, nationaler und transnationaler bzw. globaler Ebene* (Naish, 2017). Auf *lokaler Ebene* kooperieren die Vorhaben häufig mit *lokalen Regierungsstrukturen*, z.B. Kommunen, und mit *Bildungseinrichtungen* wie Schulen, Berufsschulen und Universitäten. Häufig werden hierbei SfE-Ansätze und -Methoden in bestehende Curricula integriert und/oder Zielgruppen dieser Einrichtungen ((Berufsschul-)Lehrkräfte, Dozierende, Studierende) zu SfE-Coaches geschult. *Lokale NGOs* sind z.T. an der Entwicklung von Methoden und/oder der Umsetzung von SfE-Fortbildungen beteiligt. *NGO-Netzwerkstrukturen* werden u.a. zur landesweiten Verbreitung der SfE-Ansätze und -Methoden genutzt (z.B. Brasilien). Auf der *nationalen Ebene* wird immer mit *Regierungsstrukturen*, d.h. Ministerien, die für Bildung, Sport oder Jugend zuständig sind, kooperiert. In zahlreichen Ländern sind auch *Sportverbände* in die Maßnahmen einbezogen, mit dem Ziel, SfE in Lizenzierungssysteme der Verbände zu integrieren (z.B. Indonesien) oder Methoden, die meist i.R. der Sportart Fußball entwickelt wurden, auf andere Sportarten zu übertragen (z.B. in Namibia mit Basketball, Netball, Volleyball). In diesem Zusammenhang wird seltener auch mit *lokalen Vereinsstrukturen* kooperiert (z.B. Mosambik). In einigen Partnerländern besteht eine Zusammenarbeit mit Institutionen, die zwar auch auf der transnationalen oder globalen Ebene operieren, durch die Arbeit aber auf einer nationalen oder aufgrund konkreter SfE-Maßnahmen auf lokaler Ebene mitwirken: *Multilaterale EZ-Organisationen*, wie bspw. UNHCR (United Nations High Commissioner for Refugees) i.R. der Arbeit in Flüchtlingslagern in Jordanien, *internationale Sportverbände* wie World Athletics in Uganda oder die FIFA mit dem Bau von Sportplätzen i.R. des RV S4DA sowie *internationale NGOs und Stiftungen* wie die Auma Obama Stiftung in Kenia. Die *DSHS, Beratungsfirmen, NGOs, Stiftungen* sowie *Sportverbände und -vereine aus Deutschland* sind auf lokaler, nationaler und transnationaler Ebene in zahlreiche Prozesse eingebunden, u.a. i.R. von (Mit-)Finanzierungen von Projekten, M&E-Aktivitäten, allgemeinen Beratungen zum Vorhaben, der Entwicklung von Ansätzen und Methoden sowie Fortbildungen von lokalen und nationalen Stakeholdern. In einigen Partnerländern werden zudem bereits bestehende deutsche EZ-Strukturen auf unterschiedliche Weise genutzt, bspw. wurden SfE-Maßnahmen in *andere GIZ-Programme* integriert (z.B. Sri Lanka) und inhaltlich auf

diese abgestimmt (z.B. Afghanistan, Palästinensische Gebiete), wodurch engere und losere Kooperationen entstanden sind. Die in den Partnerländern involvierten deutschen Stakeholder sind in der Abbildung in hellgrau und mit gerundeten Ecken dargestellt. Sie sollen an dieser Stelle genauer erläutert werden, um hervorzuheben, welche Bandbreite deutscher Akteur*innen an den SfE-Maßnahmen der dt. EZ beteiligt sind.

Während eine im Jahr 2012 durchgeführte Akteursanalyse ergibt, dass trotz der Wertschätzung des Sports als mögliches Instrument in der EZ, die Ansätze und Akteur*innen in Deutschland noch überschaubar und unübersichtlich sind (Doll-Tepper, Petry, Dumon, Koenen, Weinberg, Meier & Giselbrecht, 2013)⁴⁸, hält die aktualisierte Akteursstudie von 2020 einen deutlichen Zuwachs von deutschen Akteur*innen fest (Petry, 2020).⁴⁹ Die staatlichen Stakeholder, Forschungseinrichtungen, Netzwerke, Stiftungen, Sport(fach)verbände und -vereine, NGOs, soziale Entrepreneurs und Wirtschaftsunternehmen sind stark untereinander vernetzt und zeichnen sich durch vielfältige Partnerstrukturen aus (ebd.). An dieser Stelle sollen lediglich jene Akteur*innen beschrieben werden, die in einem partnerschaftlichen Verhältnis zum den Vorhaben der GIZ stehen und/oder aktiv am BMZ-Thementeam teilnehmen.⁵⁰

*Staatliche Akteur*innen*

Der bereits im Detail beschriebenen sportbezogenen EZ, die über das BMZ gesteuert wird, steht die *internationale Sportförderung des Auswärtigen Amtes (AA)* gegenüber. Jährlich werden für die Kurz- und Langzeitprojekte ca. fünf Millionen Euro zur Verfügung gestellt, 70 Prozent der Mittel fließen in Projekte in afrikanischen Ländern (AA, 2018, 07. März). Sport wird hierbei maßgeblich zur Förderung eines positiven Deutschlandbildes

⁴⁸ Es wurden 80 Akteur*innen identifiziert, davon wurden abhängig des möglichen Beitrages, den die Akteur*innen „[...] für die strategische Ausrichtung des Sports i.R. der Entwicklungszusammenarbeit leisten können“ (Doll-Tepper et al., 2013, S. V), 15 Schlüsselakteur*innen (z.B. GIZ, BMZ, DSHS, ICSSPE, street-footballworld) und 65 Sekundärakteur*innen (z.B. Kickfair, DSJ, Engagement Global, Dirk Nowitzki Stiftung) ausgemacht.

⁴⁹ Die Studie listet 100 Akteur*innen, 22 der ersten Studie wurden nicht mehr aufgenommen, dafür wurden 42 neue Institutionen identifiziert (Petry, 2020). Beide Studien wurden im Auftrag des BMZ/der GIZ durchgeführt.

⁵⁰ Eine Übersicht zu allen identifizierten deutschen SfE-Stakeholdern, inklusive ihres Beitrags zu den SDGs, gibt die Masterarbeit von Maith (2021).

im Ausland und als zentrales Element glaubwürdiger und nachhaltiger Entwicklungspolitik instrumentalisiert.⁵¹ Präziser formuliert soll Sport durch den interkulturellen Dialog zur Konfliktprävention, zur Sympathiewerbung und zur Vermittlung eines modernen Deutschlandbildes sowie zur Unterstützung der allgemeinen Ziele der deutschen Außenpolitik beitragen (Wnendt, 2011). Im Rahmen der internationalen Sportförderung findet aber auch Unterstützung im Auf- und Ausbau des Breiten- und Behindertensports als eine Art Hilfe zur Selbsthilfe statt (Groll & Hillbring, 2011). Im Zeitraum von 1961-2010 hat das AA in Kooperation mit den Hauptpartner*innen DOSB, DFB, dem Deutschen Leichtathletik-Verband (DLV) und der Universität in Leipzig mehr als 1400 Sportprojekte in 100 Ländern durchgeführt (DOSB, o.J.). Das AA und BMZ arbeiten nicht direkt zusammen; sie haben ihre eigenen Organisationsstrukturen und Konzepte (vgl. Groll & Hillbring, 2011, S. 21ff.). Seit 2013 treffen sich Vertretende der Ministerien i.R. des Thementeam SfE und tauschen sich zu ihren jeweiligen Aktivitäten aus (Petry & Bauer, 2016).

Neben den beiden staatlichen Hauptakteur*innen nimmt auch *das Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMU)* und das *Bundesministerium des Inneren (BMI)* an Sitzungen des Thementeam SfE teil. „Das BMI [übernimmt] für die Bundesregierung sportpolitische Aufgaben in internationalen Gremien; die anderen Ressorts [...] arbeiten zu“ (BMZ, 2015, S. 21). Eine weitere zugehörige Institution des Thementeam ist *Engagement Global*, die im Auftrag des BMZ hauptsächlich für entwicklungspolitische Bildungsarbeit in Deutschland zuständig ist.⁵² In anderen Ländern kooperieren die Sportvorhaben der GIZ mit anderen GIZ-Vorhaben, bspw. aus dem Bereich Bildung oder Berufsbefähigung (S4D Resource Toolkit, o.D.b).

Sport(fach)verbände und -vereine

Die unterschiedlichen Konzepte der Ministerien schlagen sich in den verschiedenen Arbeitsbereichen des DOSBs nieder. Die internationale Arbeit des Dachverbandes gliedert

⁵¹ Hierfür werden unterschiedliche Initiativen geschaffen: Bspw. die Initiative „Sport und Außenpolitik“ i.R. der Fußball-WM 2010 in Südafrika (Groll & Hillbring, 2011; Wnendt, 2011) oder die Initiative „Deutsche Fußball Botschafter e.V.“ (vgl. AA, 2018, 09. Mai).

⁵² Zum Kirchentag 2019 unterzeichnete Engagement Global mit dem DOSB eine Kooperationsvereinbarung zur gemeinsamen entwicklungspolitischen Bildungsarbeit (Engagement Global, 2019).

sich in drei Themenfelder: *Internationale Interessenvertretung*, *Internationale Sportförderung* und *Sport für Entwicklung*. Letzterer Themenbereich ist durch das 2012 unterzeichnete Partnerschaftsabkommen zwischen GIZ und DOSB mehr in den Vordergrund gerückt. Seit 2015⁵³ setzten die Partner*innen SfE in unterschiedlichen Ländern um, wobei es um die *Entwicklung durch den Sport* geht. 2017 wurde die Kooperationsvereinbarung vertieft und ausgeweitet. Ergebnisse der Unterzeichnung sind weitere gemeinsame Projekte in der Türkei, Jordanien, Kolumbien und Uganda.⁵⁴ Bei der *Internationalen Interessenvertretung* geht es um die *Entwicklung im Sport*. Hier tauscht sich der DOSB mit anderen internationalen Sportverbänden in Gremien aus und „[...] gestaltet die europäische wie internationale Sportpolitik mit“ (Grafarend & Gaerte, 2020, S. 169). Die *Internationale Sportförderung*, die über das AA finanziert wird, beschäftigt sich mit der *Entwicklung des Sports*, v.a. auch des Breitensports, in Form von Kurz- und Langzeitprojekten. Bei den mehrjährigen Langzeitprojekten steht der „[...] Auf- und Ausbaus tragfähiger Sportstrukturen auf nationaler und regionaler Ebene“ in unterschiedlichen Sportarten (v.a. Fußball, Basketball und Leichtathletik) im Vordergrund (ebd., S. 171).

Durch die Entsendung von Expert*innen ins Ausland sind der DFB und der *Deutsche Leichtathletik Verband (DLV)* schon lange in die Sportfördermaßnahmen der Bundesregierung involviert. Ähnlich der Maßnahmen an der Sportuniversität Leipzig, führt der DFB mit der Unterstützung des AA seit 1976 Trainer*innenlehrgänge an der Sportschule in Hennef durch, der DLV an der Auslandstrainerschule in Mainz (vgl. Groll & Hillbring, 2011; Steinmann, 2011). Der DFB engagiert sich seit 2013 durch die Zusammenarbeit mit dem SV SfE der GIZ in unterschiedlichen Partnerländern (Kolumbien, Brasilien, Palästinensische Gebiete), im Bereich SfE (DFB, 2016, 26. Mai).⁵⁵ Die Zusammenarbeit zwischen DFB und BMZ wurde 2016 durch ein Partnerschaftsabkommen verstärkt. Wichtiger Teil der Partnerschaft ist die Sonderinitiative „Mehr Platz für Sport – 1000 Chancen

⁵³ Ein erstes gemeinsames Projekt wurde 2015 in Namibia initiiert (vgl. Grafarend & Gaerte, 2020, S. 173; Klingmann et al., 2020, S. 147f.).

⁵⁴ Weitere Aktivitäten im Themenfeld SfE (bspw. zu Botschaftern und dem Programm *weltwärts*) s. Grafarend und Gaerte (2020, S. 174f.).

⁵⁵ „Mit der Aufnahme des Prinzips der Nachhaltigkeit in seine Satzung hat sich der DFB bereits 2010 verpflichtet, durch Fußballprojekte die soziale Entwicklung zu fördern. Auf der Basis dieses geteilten Verständnisses tragen DFB und BMZ entwicklungspolitische Themen in die Öffentlichkeit und führen gemeinsam Vorhaben in Entwicklungs- und Schwellenländern durch“ (BMZ, 2015, S. 21).

für Afrika“ (s.o.) sowie „die Integration und Förderungen von Flüchtlingen durch Fußball“ mit der Entsendung von Expert*innen (bspw. nach Afghanistan, Jordanien, Kolumbien, Nordirak und in die Palästinensischen Gebiete) (DFB, 2016, 29. Mai). Eine weitere wichtige Komponente der Zusammenarbeit ist der *Internationaler Expert*innenlehrgang SfE*, der in Deutschland bereits drei Mal durchgeführt wurde. Hierbei werden Multiplikator*innen aus den Partnerländern der GIZ-Vorhaben zu einem ca. einwöchigen Kurs zu SfE eingeladen (DFB, 2016, 26. Mai; 2019, 23. August). Zudem werden 2018 die Nationalspieler*innen Gerald Asamoah und Nia Künzer zu *Botschafter*innen Fußball für Entwicklung* benannt (BMZ, 2018, 26. März). Der DLV kooperiert i.R. des Projekts „Athletics for Development“ in Uganda seit 2018 mit der GIZ (BMZ, 2018, 11. Mai). Neben dem *TSG Hoffenheim*, mit dem die GIZ erstmalig i.R. des Projekts in Namibia zusammenarbeitete, ging das SV SfE mit *ALBA Berlin* und *1. FC Köln* weitere Vereinsk Kooperationen ein (Birnbrich, 2021; GIZ, 2020; S4D Resource Toolkit o.D.a).

Forschungseinrichtungen

Neben der DSHS ist auch die *Universität Leipzig* Teil des Thementeam. An der Universität werden für ausländische Trainer*innen zwei Mal jährlich fünfmonatige Trainer*innenlehrgänge in verschiedenen Sportarten angeboten, die seit 1991 durch das AA gefördert werden. Die Teilnehmenden sollen dazu befähigt werden, eigenständig und nachhaltig Sportstrukturen in den jeweiligen Herkunftsländern aufzubauen (vgl. Groll & Hillbring, 2011, Wnendt, 2011). Die DSHS ist seit 2013 wichtige*r Partner*in hinsichtlich der Sicherung der „Qualität und Wirksamkeit und damit [der] Nachhaltigkeit“ der Arbeit des BMZ (BMZ, 2015, S. 22; S4D Resource Toolkit, o.D.b). Die Hochschule begleitet das SV SfE der GIZ wissenschaftlich – von Mitte 2013 bis Mitte 2016 in einer ersten Phase in Kooperation mit dem *Weltrat für Sportwissenschaft und Leibes-/Körpererziehung (ICSSPE)*, Mitte 2016 bis Mitte 2019 in einer zweiten und ab Ende 2020 bis Ende 2022 in einer dritten Phase. Hierdurch hat die DSHS die Möglichkeit, das SV SfE von seinen Anfängen bis heute zu beraten und erhält umfangreiche Einblicke in alle bisherigen Phasen. Auch das RV S4DA wurde zeitweise wissenschaftlich begleitet (vgl. Schmid & Arnemann,

2020, S. 161), im RV Jordanien und Irak berät die DSHS auf Anfrage über das SV SfE, im Vorhaben in der Türkei war die DSHS über den DOSB involviert (vgl. Neeten, 2020).⁵⁶

Netzwerke

Das globale *Netzwerk von ICSSPE* besteht aus ca. 250 nationalen wie internationalen Sport- und Sportwissenschaftsinstitutionen und Sportministerien. Es hat zum Ziel, die Zusammenarbeit zwischen Wissenschaftler*innen, politischen Entscheidungsträger*innen und Praktiker*innen zu fördern. Über zahlreiche Veröffentlichungen strebt ICSSPE die Vermittlung von sportwissenschaftlichem und sportpädagogischem Hintergrundwissen für Akteur*innen im Bereich SfE an (Petry, 2020). In der ersten Phase des SV SfE war ICSSPE an der wissenschaftlichen Begleitung beteiligt. Das NGO Netzwerk *streetfootball-world (sfw)* ist aufgrund seiner langjährigen Expertise im SfE-Kontext ebenfalls Teil des Thementeam und umfasst über 100 Mitglieder in zahlreichen Ländern. Das Netzwerk hat erstmalig in Brasilien umfangreicher mit der GIZ kooperiert (S4D Resource Toolkit, o.D.b). Die *Methode Football3* kommt in zahlreichen Partnerländern zum Einsatz.⁵⁷

NGOs und Stiftungen

Mit *CARE Deutschland e.V.* initiierte die GIZ bereits sportbezogene Bildungsprojekte in Kenia und Bosnien, mit *Don Bosco e.V.* in Kenia und Äthiopien. Die *TUI Care Foundation* war als Teil des Nachhaltigkeitsnetzwerks *Futouris* in GIZ-Projektaktivitäten in Namibia involviert (Petry, 2020; S4D Resource Toolkit. O.D.b). Die deutschen Vertretenden der internationalen NGO *Right to Play* sowie der *Senior Experten Service (SES) - Stiftung der Deutschen Wirtschaft für internationale Zusammenarbeit* sind aufgrund ihrer Expertise im SfE-Kontext und in der EZ Teil des Thementeam SfE in Deutschland.

*Wirtschaftsunternehmen und Soziale Entrepreneurs*⁵⁸

Seit 2012 ist das Interesse von Wirtschaftsunternehmen am SfE-Thema stark gestiegen. Sie befassen sich schwerpunktmäßig mit der nachhaltigen Produktion von Sportartikeln

⁵⁶ Aufgaben und Arbeitsbereiche der DSHS werden in Kap. 6 detailliert beschrieben.

⁵⁷ S. Ausführungen unter Kapitel 2.3.3.

⁵⁸ Darunter werden ‚soziale Unternehmen‘ verstanden, die „gesellschaftliche Verantwortung übernehmen“ und „zur Lösung oder Verbesserung gesellschaftlicher Missstände“ beitragen wollen (Achleitner, 2018, 14. Februar).

und fairem Handel (Petry, 2020). In Südafrika und Brasilien kooperiert die GIZ bereits 2012 mit *Nike Inc.* i.R. der Kampagne *Design to Move*, mit dem Fokus auf der Förderung der physischen Gesundheit von Kindern und Jugendlichen im Schulkontext (S4D Resource Toolkit, o.J.b). Deutsche Unternehmen wie *VAUDE* oder *Jako AG* sind Teil des Textilbündnisses (Petry, 2020), eine offizielle Kooperation zum SV SfE steht noch aus. Durch das Unternehmen *SPIN Sports Innovation*, das von einem ehemaligen DSHS-Mitarbeiter geleitet wird, wurden zahlreiche SfE-Ansätze und -Methoden mitgestaltet und entwickelt, bspw. in den Palästinensischen Gebieten, Indonesien und in der Region Westbalkan (vgl. bspw. Steinbach & O'Rourke, 2017). Mit der Beratungsfirma *GOPA* arbeitet das RV S4DA in Äthiopien zum Thema berufliche Bildung (*GOPA*, o.D.). Gemeinsam mit dem in Hamburg ansässigen Unternehmen *Viva con Agua* wurde in Uganda ein Hygiene-Handbuch für Kinder mit Corona-Verhaltensregeln erstellt und an zahlreichen Schulen verteilt. Das soziale Beratungsunternehmen *CampGroup* unterstützte das SV SfE bei der Entwicklung des SfE-Ansatzes und -Methode in Afghanistan bereits 2015 (CampGroup, 2015, 28. Oktober; GIZ SV SfE, 2021).

Abbildung 3 fasst alle beschriebenen Akteur*innen, die an SfE-Maßnahmen der dt. EZ bereits mitwirkten, zusammen. Mitwirkende im BMZ-Thementeam SfE sind hellgrau gekennzeichnet. Dunkelgrau markierte Akteur*innen sind nicht Teil des Thementeam, kooperieren oder kooperierten jedoch bereits mit der GIZ i.R. unterschiedlicher SfE-(Teil-)Maßnahmen, so wie auch alle fett gekennzeichneten.

2 Hintergrund

Staatliche Akteur*innen	GIZ	BMZ	AA	BMU	BMI	Engagement Global
Sportverbände & -vereine	DOSB	DFB	DLV	TSG Hoffenheim	ALBA Berlin	1. FC Köln
Forschungseinrichtungen	DSHS	Universität Leipzig				
Netzwerke	ICSSPE	streetfootballworld				
NGOs & Stiftungen	CARE Deutschland	Don Bosco Mondo	Futouris	Right to Play	TUI Care	SES, Stiftung der Dt. Wirtschaft für internationale Zusammenarbeit
Soziale Entrepreneurs & Wirtschaftsunternehmen	Camp Group	Viva con Agua	SPIN	GOPA-		

Abb. 3: Deutsche SfE-Stakeholder (eigene Darstellung)

2.3 Terminologische Reflexion: Der Versuch einer begriffsbestimmenden Einordnung von ‚Sport für Entwicklung‘ bzw. von Sport im deutschen EZ-Kontext

Anhand des sich veränderten Gebrauchs von Begriffen in der Sportberichtserstattung der Bundesregierung und betreffender Akteur*innen, der Entwicklung definitorischer Modelle sowie sportwissenschaftlicher und pädagogischer Erklärungsmodelle des Sports, erfolgt im Folgenden eine begriffsbestimmende Einordnung von *Sport* und *Entwicklung*.

Daraus resultierend kann eine präzisere Antwort auf die Frage, von *welchem Sport* im deutschen SfE-Kontext gesprochen wird, gegeben werden. Durch die genauere Betrachtung der sportlichen Inhalte der SfE-Maßnahmen soll ein Beitrag geleistet werden, die Definition von SfE für den deutschen Kontext zu spezifizieren. Eine allgemeingültige und damit internationale Einordnung und Präzisierung kann in dieser Arbeit nicht geleistet werden, wäre allerdings ein wichtiger Beitrag zum Forschungsfeld SfE und demnach als Gegenstand für weitere Forschungsvorhaben zu empfehlen.

2.3.1 Von der Sportförderung zu ‚Sport für Entwicklung‘?

Sportförderung in Ländern der Dritten Welt, Förderung des Sports in den Entwicklungsländern, Sportentwicklungshilfe, Sportförderung i.R. der Entwicklungspolitik, Entwicklungszusammenarbeit auf dem Gebiet des Sports, Sport in der Entwicklungszusammenarbeit, sportbezogene Entwicklungszusammenarbeit, Entwicklungszusammenarbeit durch Sport, Sport und Entwicklung(-zusammenarbeit), Sport für Entwicklung – dies seien nur einige Umschreibungen eines politischen Handlungsfeldes, das sich im Laufe der Zeit zunehmend ausdifferenziert hat (vgl. Groll & Hillbring, 2011).

Auch wenn die GTZ bereits in den 1980er Jahren auf die präzise Unterscheidung der Sportfördermaßnahmen in den Entwicklungsländern verwies (Dohrmann et al., 1985),⁵⁹ war es noch ein langer Weg bis zu einer präziseren und gebräuchlicheren Unterscheidung und der Etablierung des Begriffs *Sport für Entwicklung*. Dies hat zum einen damit

⁵⁹ Es wurde unterschieden zwischen „der Promotion von Sport in Entwicklungsländern, im Wesentlichen als Public-Relation-Maßnahme [...], der Förderung/der Entwicklung des Sports in Entwicklungsländern, d.h. auf den Auf- und Ausbau des Sports gerichtete Maßnahmen, der Entwicklungszusammenarbeit über den Sport bzw. mit dem Sport, d.h. der Einsatz von Sport für den Gesamtkomplex der EZ in den Entwicklungsländern“ (Dohrmann et al., 1985, S. 16f.).

zu tun, dass sich in Deutschland im Laufe der Zeit die Maßnahmen der Ministerien verändert bzw. ausdifferenziert haben: Ist der Auftrag des AA im Bereich des Sports seit 1960er Jahren derselbe, nämlich die Förderung des Sports i.R. der auswärtigen Kulturpolitik,⁶⁰ so hat sich der Auftrag des BMZ nach den sich wandelnden entwicklungspolitischen Diskussionen und Agenden gerichtet (bspw. in den 90er Jahren, MDGs, SDGs; vgl. bspw. BMZ, 2018). Zum anderen haben sich die Schwerpunkte der Durchführungsorganisation im Laufe der Zeit verschoben oder auch ausdifferenziert: Während die GTZ einst im Auftrag des AA zahlreiche Sportfördermaßnahmen realisierte (vgl. Kap. 2.1), setzt die GIZ heute ausschließlich Sport im Auftrag des BMZ als Instrument in ihrer entwicklungspolitischen Arbeit ein. Sportverbände (bspw. DOSB, DFB, DLV) setzten früher mehrheitlich Projekte zur Sportförderung um und entdeckten bzw. intensivierten den ‚neuen Themenbereich‘ erst spät (vgl. Grafarend & Gaerte, S. 173).

Die Entwicklung von einer ‚reinen Sportförderung‘ in Ländern des Globalen Südens hin zu einer zusätzlichen Instrumentalisierung des Sports zur Erreichung entwicklungspolitischer Zielsetzungen, und somit auch die organisatorische und inhaltliche Entwicklung des Einsatzes von Sport i.R. von Entwicklungsmaßnahmen, lässt sich für den deutschen Sprachgebrauch anhand der *Sportberichtserstattung der Bundesregierung* nachvollziehen: Bereits Anfang der 1980er Jahre wird im fünften Sportbericht der Bundesregierung zum ersten Mal von den „Entwicklungspolitischen Impulsen des Sports“ gesprochen (Dt. Bundestag, 1982, S. 69).⁶¹ Die Maßnahmen des AA und BMZ werden zwischen Mitte der 1970er bis Ende der 1980er unter gemeinsame Überschriften gefasst,⁶² auf die ab den 1990er bis Anfang der Nullerjahre verzichtet wird (vgl. Dt. Bundestag, 1990, 1995, 1999, 2002)⁶³. Erst ab dem elften Sportbericht (Dt. Bundestag, 2006) und bis heute stehen die

⁶⁰ S. hierzu auch die Beschreibungen in den Sportberichten der Bundesregierung.

⁶¹ „Vom Sport können aufgrund seiner pädagogischen und sozialen Auswirkungen erhebliche Impulse auf den Entwicklungsprozess in den Ländern der Dritten Welt ausgehen. Dies äußert sich nicht nur durch die Übertragung des im Sport eingeübten Verhaltens von Individuen und Gruppen auf andere gesellschaftliche Bereiche; vielmehr vermag der Sport darüber hinaus Kommunikationsprozesse zwischen verschiedenen ethnischen und sozialen Gruppen herbeizuführen, soziale Schranken abzubauen und auf diese Weise gemeinschaftsbildend zu wirken“ (Dt. Bundestag, 1982, S. 69).

⁶² Bspw. als „Förderung des Sports in den Entwicklungsländern“ ab dem dritten Sportbericht (vgl. Dt. Bundestag, 1976), „Internationale Sportpolitik, Sportentwicklungshilfe“ ab dem fünften Sportbericht (vgl. Dt. Bundestag, 1982)

⁶³ Als getrennte Kapitel: „Sportförderung im Rahmen der Auswärtigen Kulturpolitik“ und „Sportförderung im Rahmen der Entwicklungspolitik“.

Maßnahmen der Ministerien unter dem Titel „Internationale Sportförderung“, was sich hauptsächlich darauf bezieht, dass die Bunderegierung alle von ihr international geförderten bzw. finanzierten Maßnahmen mit Sportbezug unter diesem Titel vereinen will, dieser sich aber von anderen Definitionen unterscheidet (vgl. Grafarend & Gaerte, 2020, S. 170ff.). Mit dem 13. Sportbericht (Dt. Bundestag, 2014) werden die BMZ-Maßnahmen erstmalig unter dem Kapitel „Sport und Entwicklung“ gefasst, die Zuständigkeit des BMZ für den Sport wird als „Sport als Instrument zur Erreichung von Entwicklungszielen“ beschrieben (ebd., S. 19). Im 14. Sportbericht (Dt. Bundestag, 2019) stehen die Maßnahmen letztendlich unter dem Kapitel „Sport für Entwicklung“. Im Rahmen der *entwicklungspolitischen Berichtserstattung* der Bundesregierung wird der Sport erstmalig neben anderen Aktionsräumen der damaligen Afrikainitiative unter Minister Dirk Niebel erwähnt (vgl. BMZ, Mai 2013, S. 115). Mit der Begründung, „um Jugendliche dort abzuholen, wo sie ihre Freizeit verbringen“ (BMZ, März 2017, S. 164), wird im 15. entwicklungspolitischen Bericht das Engagement im Bereich *Sport und Entwicklung* beschrieben, wobei lediglich auf die Baumaßnahmen i.R. der Initiative „Mehr Platz für Sport – 1000 Chance für Afrika“ verwiesen wird.

Der Wandel hinsichtlich der Kombination der Begriffe *Sport* und *Entwicklung* ist auch von *internationalen Entwicklungen* geprägt: Die UN stehen seit Anfang der 2000er für den Begriff *Sport for Development and Peace (SDP)*.⁶⁴ Ein zweidimensionales Modell, *plus sport - sport plus* (vgl. Coalter & Taylor, 2010, S. 1) oder auch *Entwicklung plus Sport – Sport plus Entwicklung* (vgl. DEZA, 2005, S. 19f.) beeinflusst ab Mitte der Nullerjahre die inhaltliche Weiterentwicklung und ist auch heute noch beliebter Diskussionsgegenstand. Levermore und Beacom (2009) wählen den Begriff *sport-in-development* „[...] as representative of the perception that the use of sport may assist the international development process“ (ebd., S. 9) . Coalter und Taylor (2010, S. 1) nehmen die Akteur*innen als Ausgangspunkte ihrer Kategorien, die „social development organizations“ für die Kategorie *plus sport* sowie Organisationen mit Sport als „core activity“ (bspw. Verbände,

⁶⁴ Die Abkürzung *SDP* wird in englischsprachiger Literatur oft verwendet und in dieser Arbeit bei direkten Zitaten beibehalten.

Vereine) für die Kategorie *sport plus*.⁶⁵ Die *Schweizer Direktion für Entwicklung und Zusammenarbeit* (DEZA, 2005, S. 19f.) wählt zur Unterscheidung ihrer Kategorien ebenfalls die unterschiedlichen Akteur*innen und charakterisiert die Stoßrichtungen wie folgt:

- „*Entwicklung plus Sport*: Ausgangspunkte für diese Projekte sind meist Entwicklungsorganisationen, Regierungen und NGOs. Sie setzen Sport als Instrument für die Erreichung spezifischer Entwicklungsziele ein, etwa für die Bekämpfung von HIV/Aids oder die Wiedereingliederung von Straßenkindern usw. Idealerweise sind diese Projekte in übergreifende Entwicklungsprogramme integriert.“
- „*Sport plus Entwicklung*: Typischerweise werden diese Projekte von Akteur*innen aus der Sportwelt und von Regierungen initiiert. Sie zielen auf die Förderung des Sports und der schulischen Sporterziehung ab und betrachten den Sport als Wert an sich. Bewusst oder unbewusst können sie aber auch zur Entwicklung beitragen, etwa durch die Bildung von gemeinschaftlichen Strukturen, die Förderung von Einkommensmöglichkeiten oder die Bereitstellung von Infrastrukturen.“

Hinsichtlich der letzteren Stoßrichtung verweist Rockenfeller (2009) auf die Nähe zum Begriff der *Sportförderung*, bei dem es hauptsächlich um den „Auf- und Ausbau tragfähiger Sportstrukturen“ (Grafarend & Gaerte, 2020) geht und um Maßnahmen, die auf „die Entwicklung des Sports abzielen, ohne einen bewussten Entwicklungsbeitrag zu leisten. Über den Aspekt, durch diese Maßnahmen den reinen Zugang zum Sport zu ermöglichen, ist kein wirklicher Entwicklungsbeitrag zu identifizieren“ (Rockenfeller, 2009, S. 5). Im Hinblick auf die unterschiedlichen deutschen Akteur*innen, die sich entweder eher dem Sport- oder eher dem Entwicklungs-Begriff zugehörig fühlen, und damit verschiedene Motive haben, ist eine umfassende Abbildung gefragt, die alle Maßnahmen, die Sport zur Erreichung von Entwicklungszielen nutzen, umfasst und zeitgleich eine

⁶⁵ Coalter und Taylor (2010, S. 1) sehen den Unterschied zwischen diesen zwei Wortkombinationen wie folgt: „Plus sport organisations being social development organisations dealing with issues such as conflict resolution, homelessness and children at risk. Sport was either part of their programme, or they were encouraged to introduce it as part of this project. Sport plus organisations refers to those whose core activity was sport, which is used and adapted in various ways to achieve certain ‘development’ objectives, such as HIV and AIDS education or female ‘empowerment’.“

klare Trennung zur Sportförderung vollzieht.⁶⁶ Das folgende Modell (Abb. 4) unterscheidet zwischen den Kategorien *Sportförderung*, womit die *Entwicklung des Sports* gemeint ist, und *Sport für Entwicklung*, die auf die *Entwicklung durch Sport* abzielt.

Entwicklung des Sports – (Int.) Sportförderung	Entwicklung durch Sport – Sport für Entwicklung (SfE)	
	Sport plus Entwicklung	Entwicklung plus Sport
Auf- und Ausbau von Sportstrukturen (bspw. Beratung von Verbänden, Trainer*innenausbildungen) und Sportinfrastrukturen; Entwicklung des Breitensports (Verbesserung des Zugangs zum Sport); die Entwicklung sportsspezifischer Kompetenzen (u.a. Technik, Taktik), Wettkämpfe.	Fokus: Sport; Akteur*innen aus dem Sport nutzen Entwicklungsziele i.R. ihrer Arbeit und ihren sportlichen Angeboten, d.h. Trainingseinheiten und Sportaktivitäten werden im Sinne von Entwicklungszielen (bspw. zur HIV/Aids-Prävention) adaptiert, modifiziert.	Fokus: Entwicklungsziele; Akteur*innen aus der EZ nutzen Sport als Zugangsmittel (bspw. zu best. Zielgruppen), als Instrument oder zusätzliche Komponente, um bestimmte Entwicklungsziele zu fördern (z.B. Berufsbefähigung).

Abb. 4: Sportförderung und Sport für Entwicklung (eigene Darstellung in Anlehnung an Coalter & Taylor, 2010; DEZA, 2005; Grafarend & Gaerte, 2020⁶⁷)

2.3.2 Von welchem ‚Sport‘ reden wir überhaupt?

Dieser Frage ein weiteres Kapitel zu widmen, scheint vor dem Hintergrund berechtigt, dass sich das Verständnis von Sport seit den Anfängen der Sportförderung in Deutschland bis heute ungemein gewandelt und ausdifferenziert hat. Zudem sind, wie in Kapitel 2.2.2 beschrieben, sehr viele Akteur*innen am Arbeitsfeld beteiligt, die den Sport aus unterschiedlichen Perspektiven betrachten. Hier hat sich nicht nur das Sportverständnis aus Sicht der Sportakteur*innen, sondern auch das Verständnis der EZ- und Partnerorganisationen, wie bspw. im Bereich des Schulsports, gewandelt.

Aktuelle Sportdefinitionen im EZ-Kontext fassen unter *Sport* zahlreiche Termini. So versteht das BMZ (2015, S. 7) in Anlehnung an eine Definition der Vereinten Nationen⁶⁸ unter Sport „[...] alle physischen Aktivitäten, die das körperliche und geistige Wohlbefinden und das soziale Miteinander fördern, vom Breiten- und Freizeitsport über Spiel und

⁶⁶ Nach der Aufschlüsselung der Begrifflichkeiten in dieser Arbeit werden die Maßnahmen der Sportförderung und der sportbezogenen EZ seitens der Bundesregierung ‚fälschlicherweise‘ alle unter der Überschrift *Internationale Sportförderung* gefasst (vgl. Dt. Bundestag, 2019).

⁶⁷ Grafarend und Gaerte (2020) haben die Begrifflichkeiten aus Akteursperspektive (DOSB) neu sortiert.

⁶⁸ “Incorporated into the definition of ‘sport’ are all forms of physical activity that contribute to physical fitness, mental well-being and social interaction. These include: play; recreation; organized, casual or competitive sport; and indigenous sports or games” (United Nations Inter-Agency Task Force on Sport for Development and Peace, 2003, S. V).

Bewegung bis hin zu traditionellen Kultur- und Ausdrucksformen wie beispielsweise Tanz“. In aktuellen internationalen Policies und Publikationen wird Sport als Oberbegriff verwendet, der *Sport for all*, Spiel und Bewegung, Freizeitsport, Tanz sowie organisierte und unorganisierte, traditionelle und indigene Sportarten und Spiele in ihrer Vielfalt umfasst (vgl. Commonwealth Secretariat, 2019, S. 8; UNESCO, 2017, S. 9).⁶⁹ Aus sportwissenschaftlicher Perspektive lohnt es sich, die hier genutzten Begriffe wie bspw. *Breitensport*, *Freizeitsport* oder *Sport für alle* genauer zu betrachten, um „alle physischen Aktivitäten“ (ebd., S.7) im Bereich SfE genauer zu verstehen und im Anschluss präziser verorten zu können.

Erklärungsmodelle des Sports

In den Sportwissenschaften gibt es umfangreiche Versuche, den Begriff Sport zu definieren (vgl. Röthig & Prohl, 2003, S. 493). Eindeutig ist, dass sich das Begriffsverständnis mit dem Wandel in der Gesellschaft, der sowohl eine Veränderung der Bewegungskultur (vgl. Größing, 1993) als auch die demographischen Veränderungen und den damit verbundenen Wertewandel beinhaltet (vgl. Blessing-Kapelke, Klages & Schwind-Gick, 2007; Steinbach, Groll, Haasner & Damerius, 2008), verändert hat. Damit einher geht der Wandel von einem *engen Sportverständnis*, bei dem die Sportarten mit eindeutig messbaren Leistungszielen, einem internationalen Regelwerk und einer Organisierbarkeit durch Wettkämpfe im Mittelpunkt stehen, hin zu einem *weiten Sportverständnis*, bei dem das Lösen von Bewegungsaufgaben mit *unterschiedlichen Zielsetzungen* verbunden ist. Die damit verbundene „Vielfalt in den Praxisformen und Sinnorientierung“ ist ein Grund für die Entstehung zahlreicher Wortverbindungen, wie bspw. Freizeitsport, Erlebnissport und Gesundheitssport (Wopp & Dieckert, 2002, S. 13). Die Ausdifferenzierung des Sportverständnisses schlägt sich in den theoretischen Modellen in der Sportwissenschaft nieder: Das ursprüngliche *Pyramidenmodell*, bei dem der Leistungs- und Spitzensport die Spitze und der *Breitensport* die breite Basis symbolisiert, wurde in den 1980er Jahren von der *Zweisäulentheorie* als dualem Modell abgelöst, die zwischen dem *sportlichen*

⁶⁹ Originalzitat: “A generic term comprising sport for all, physical play, recreation, dance and organised, casual, competitive, traditional and indigenous sports and games in their diverse forms.”

und *rekreativen Prinzip* unterscheidet, und den Begriff des *Freizeitsports* in den Mittelpunkt rückt. Im weiteren Verlauf entstanden zahlreiche *plurale Erklärungsmodelle*, die versuchten, die Vielfalt in Motiven und Praxisformen des Sporttreibens abzubilden (Wopp & Dieckert, 2002). Wopp (1995) setzte mit seinem Modell einen vorläufigen Schlusspunkt in der Diskussion. In seinem dualen Modell unterscheidet er zwischen einer *leistungssportlichen und freizeitsportlichen Ausprägung des Sports*, wobei er letztere in eine breiten-sportliche Ergebnisorientierung (*Breitensport*), eine soziokulturelle Erfahrungsorientierung (*Erlebnissport*) und alternative Erfahrungsorientierung (*Alternativ-sport*) differenziert. Diese Begriffe sind wiederum Bezeichnungen für weitere Sammelbegriffe wie bspw. Fitnes-sport, Gesundheit-sport, Abenteuer-sport unter dem Begriff Erlebnissport. Der Begriff Freizeitsport wird hierbei „als Programmformel zur Charakterisierung eines Sporttreibens möglichst vieler Menschen (Sport für alle) verwendet“ (Wopp & Dieckert, 2002).

Der Begriff *Sport für alle* ist ursprünglich durch den DSB geprägt, der mit dem 1959 beschlossenen „zweiten Weg“ und den darauffolgenden *Trimm-dich-Kampagnen* das Sporttreiben möglichst allen Menschen, v.a. denen, die bislang davon ausgeschlossen waren, zugänglich machen wollte (Bauer, 2004). Er wird z.T. mit dem Begriff Breiten-sport synonym verwendet (vgl. Trosien, 1987, S. 124f.).⁷⁰ „Der Begriff ‚Sport für alle‘ hat mittlerweile unter dem Slogan ‚Sport for All‘ Eingang in den internationalen Sprachgebrauch und in internationale Vereinbarungen gefunden und hat sich zu einem weltweit genutzten Terminus entwickelt“ (Bauer, 2004, S. 56), bspw. als „Idee der universellen Inklusion der Bevölkerung in den Sport [...]“ (Hartmann-Tews, 1996, S. 40).

⁷⁰ Für Bauer (2004, S. 58) hebt „die Bezeichnung Breiten-sport die durch den Sport angesprochene Zielgruppe hervor, nämlich die Menschen, die die breite Basis der Pyramide bilden als die Basis des Pyramidenmodells [...] während ‚Sport für alle‘ für das sportpolitische Programm des DSB mit dem Ziel, möglichst vielen Menschen die Möglichkeit und Gelegenheit zum Sporttreiben zu geben, steht“.

Die Pädagogik kommt ‚ins Spiel‘

Allgemein betrachtet, sieht Größing (1993) die *Spielfähigkeit* als übergeordnete pädagogische Zielsetzung sportlicher Aktivitäten: Durch Spielsituationen in Gruppen (im inner- und außerschulischen Kontext) werden „soziale Erfahrungen über das Zusammenleben und die Zusammenarbeit“ gemacht, die bspw. zur Entwicklung von Regelbewusstsein und sozialer Fähigkeiten (z.B. Teamarbeit, Kommunikation, Kooperation) beitragen und in andere Lebensbereiche übertragen werden können (vgl. ebd., S. 48f.).⁷¹

Um zu verstehen, welche pädagogischen Aspekte bei SfE-Maßnahmen und -Aktivitäten eine Rolle spielen, müssen allerdings unterschiedliche Teildisziplinen der Pädagogik genauer betrachtet werden. Durch die Kombination mit dem Begriff der Pädagogik sind zahlreiche neue Handlungsfelder oder Teildisziplinen⁷² entstanden, die alle auch eine sport- und bewegungsbezogene Relevanz haben: *Sportpädagogik*, *Freizeitpädagogik/Pädagogik der Freizeit*⁷³ und *Erlebnispädagogik*.⁷⁴

Kurz gesagt, beschäftigt sich die *Sportpädagogik* mit der *Bildung und Erziehung im und durch Sport*. In ihrem Blickpunkt liegen Bildungseinrichtungen (Schulen, Kindergärten), aber auch außerschulische Sport- und Bewegungssettings (bspw. Vereine, Freizeit etc.) (Balz & Kuhlmann, 2009).⁷⁵ An Schulen haben die seit ca. 2011 eingeführten *kompetenzorientierten Lehrpläne* und damit verbundene pädagogische Leitideen Einfluss auf den Schulsport (Neuhaus, Barthel, Goeke, Joppien, Kather, Nitsche, Plener, Schlecht & Stach, 2016). Der Auftrag des *Schulsports* zur einfachen ‚Leibeserziehung‘ hat sich zu einem

⁷¹ Mit der Disziplin verbunden sind Wortzusammensetzungen, wie *Bewegungspädagogik*, *Sportpädagogik* und *Freizeitpädagogik*. Zum Verhältnis zwischen der Freizeitpädagogik und Sportpädagogik s. Michels, Petry und Tokarski (2003, S. 156ff.).

⁷² Sport-, Freizeit- und Erlebnispädagogik werden hier als Teildisziplinen der Pädagogik verstanden. Zur Ausdifferenzierung der Teildisziplinen für die Pädagogik s. Kron, Jürgens und Standop (2013, S. 24ff.).

⁷³ Ende der 90er Jahre wurde ein neues Selbstverständnis der pädagogischen Freizeitforschung formuliert, das Auswirkungen auf die verwendeten Begrifflichkeiten hatte. S. hierzu Popp und Schwab (2003, S. 14ff.) und Opaschowski (2003, S. 15).

⁷⁴ Die Verhältnisse zwischen den Teildisziplinen erörtern bspw. Michels et al., (2003) für die Freizeitpädagogik und Sportpädagogik und Fischer (2003) für die Freizeitpädagogik und Erlebnispädagogik.

⁷⁵ Balz und Kuhlmann (2009, S. 27) verorten die Sportpädagogik als Schnittstelle zwischen den Erziehungswissenschaften und den Sportwissenschaften. Zur Charakterisierung und ausführlichen Definitionen rund um die Sportpädagogik sowie zur Begrifflichkeit *Leibeserziehung* in diesem Zusammenhang Wydra (2008) und Mittmannsgruber (2005).

Doppelauftrag entwickelt: Neben der Erschließung der Bewegungs-, Spiel und Sportkultur oder der *Erziehung zum Sport*, geht es auch um die Entwicklungsförderung durch Bewegung, Spiel und Sport, also der *Erziehung durch Sport* (vgl. Kurz, 2008; Ministerium für Schule und Weiterbildung des Landes NRW, 2014; Zeuner, 2015).⁷⁶

Michels, Petry und Tokarski (2003) nehmen in diesem Zusammenhang im Hinblick auf den Freizeitsport Gebrauch des Begriffs des *Lernens*, worauf in pädagogischen Situationen gezielt Einfluss genommen wird. Sie unterscheiden zwischen *Lernen für den Freizeitsport*, womit bspw. das Lernziel ‚Lebenslanges Sporttreiben‘ gemeint ist, sowie *Lernen im* und *Lernen durch* Freizeitsport. Bei der Dimension *Lernen im Freizeitsport* stehen Lernprozesse im Zentrum, die das Mitmachen im Freizeitsport ermöglichen, d.h. alle motorischen und sozialen Fähigkeiten, die hierfür benötigt werden. Die Dimension *Lernen durch den Freizeitsport* beinhaltet den Erwerb von Kompetenzen (bspw. soziale Kompetenzen) für Lebensbereiche, die über den Freizeitsport hinausgehen, d.h. es findet eine Transferleistung des Lernens in den Alltag statt.⁷⁷

Bei der *Erlebnispädagogik* steht das Lernen durch Erlebnisse im Vordergrund. Nur eine von vielen Definitionen ist jene von Heckmair und Michl (2008, S. 75), die der Ansicht sind, dass „Erlebnispädagogik [...] eine handlungsorientierte Methode [ist], [die] durch exemplarische Lernprozesse, in denen junge Menschen vor physische, psychische und soziale Herausforderungen gestellt werden, diese in ihrer Persönlichkeitsentwicklung fördern und dazu befähigen, ihre Lebenswelt verantwortlich zu gestalten“.⁷⁸ Auch wenn sich u.a. die übergeordneten Ziele, die didaktische Gestaltung, Methoden und Arbeitsebenen im „erlebnispädagogischen Grundwortschatz“ unterscheiden, so zählen zu den Formen der Aktivitäten hauptsächlich *sportbezogene Aktivitäten* (vgl. Baig-Schneider, 2012, S. 19ff.), was zahlreiche Spielsammlungen und Praxisbeispiele verdeutlichen (vgl.

⁷⁶ Zur Kritik und Herausforderungen der neuen Lehrpläne im Schulsport s. Kurz (2008) und Zeuner (2015).

⁷⁷ Auch in diesem Bereich gehört „der Nachweis solcher Transferleistungen des Lernens [...] zu den schwierigen und bisher kaum empirisch belegten Forschungsfeldern des Freizeitsports“ (Michels et al., 2003, S. 161).

⁷⁸ Da in der Erlebnispädagogik zahlreiche Ansätze und Theorien anzutreffen sind, gibt es viele Definitionen. S. hierzu Baig-Schneider (2012, S. 13ff.). Baig-Schneider (2012) erörtert auch ob und inwiefern Erlebnispädagogik eine (Teil-)Wissenschaft, ein Verfahren oder eine Methode ist (ebd., S. 22ff.) und gibt einen Überblick zu fundamentalen Ansätzen, wie der Erlebnistherapie nach Kurt Hahn oder der Handlungspädagogik nach John Dewey (ebd., S. 27ff.). Zur Entstehungsgeschichte und Institutionalisierung durch Hahn s. auch Michl (2009).

Böhnke, 2010; Reiners, 2007, 2014). Ob die Aktivitäten dem Outdoor-, Natur-, Erlebnis- oder Abenteuersport zuzuordnen sind, bleibt an dieser Stelle nebensächlich (vgl. Baig-Schneider, 2012, S. 197, 168; Böhnke, 2010; Michl, 2009, S. 81). Interessant ist, dass sich zahlreiche Kapitel des *Handbuchs Erlebnispädagogik* (Michl & Seidel, 2018) mit thematischen Inhalten und Zielsetzungen von SfE decken: *Bildung und Kompetenzerwerb, Reflexion, Inklusion, Bildung für nachhaltige Entwicklung, Migration, Gender und Diversity*. Zudem sind die von Reiners (2007) als *Interaktionsspiele*⁷⁹ bezeichneten Aktivitäten in zahlreichen SfE-Manualen/Handbüchern zu finden (vgl. GIZ SV SfE, Juli 2016; Muth & Wirsich, 2019; Steinbach & O'Rourke, 2017). U.a. durch Interaktionsspiele werden in der Erlebnispädagogik *Ereignisse* kreiert, die durch die Verarbeitung der Teilnehmenden zu *Erlebnissen* werden. Um aus den Erlebnissen einen Lerneffekt zu machen, bedarf es ausgewählter und differenzierter *Reflexionsmethoden*, die von Pädagog*innen angeleitet werden.⁸⁰ Bei der Reflexion werden Erlebnisse mitgeteilt und diskutiert, sie etablieren sich zu *Erfahrungen*. Wenn aus Erfahrungen Wissen generiert wird (*Konstruktion*), das zusätzlich durch pädagogische Expert*innen ergänzt wird (*Instruktion*), entstehen *Erkenntnisse*, die in den Alltag transferiert werden können („Was habe ich gelernt? Was kann ich hiervon für meinen Alltag nutzen?“) (Michl, 2009). Michl (2009, S. 10) ist der Ansicht, dass „wir [...] erst dann von Erlebnispädagogik [sprechen], wenn nachhaltig versucht wird, die Erlebnisse durch *Reflexion und Transfer* pädagogisch nutzbar zu machen“.⁸¹ Die Frage nach dem Transfer oder nach der *Wirksamkeit* von erlebnispädagogischen Maßnahmen oder einfacher ausgedrückt, ob und inwiefern Teilnehmende ihre Lernerfahrungen in Alltagssituationen übertragen, bleibt auch in der Erlebnispädagogik schwer zu beantworten, da die „[...] Wirkungen nur selten linear und kausal auf das pädagogische Handeln allein zurückzuführen sind“ (Reiners, 2007, S. 19, vgl. Michl, 2009, S. 50ff.). Nach Michl (2009) sind *gut gewählte Reflexionsmethoden entscheidend*, um

⁷⁹ Die Interaktionsspiele werden hier als Technik der *Interaktionspädagogik* verstanden, die in den komplexen Begriffsbereich des *sozialen Lernens* fällt, den Fokus auf das soziale Verhalten von Lernenden legt und zum Ziel hat, zwischenmenschliches Verhalten zu ändern und zu bessern (Reiners, 2007). Die Interaktionspädagogik nach George Mead als ein neuer Ansatz ergänzt die Ansätze nach Hahn und Dewey. Gemeinsam machen sie die moderne Erlebnispädagogik aus (vgl. Baig-Schneider, 2012, S. 119ff., 155ff.).

⁸⁰ Ausgewählte Methoden s. Michl (2009, S. 51ff.).

⁸¹ Nach Michl (2009, S. 9f.) müssen sich die Ereignisse und die Reflexion i.R. der Erlebnispädagogik die Waage halten: „Werden nur Erlebnisse angeboten, [...] befinden [wir] uns im breiten Feld der Freizeitpädagogik. Befassen wir uns hauptsächlich mit der Auswertung von Erlebnissen [...], so haben wir es eher mit dem Bereich der Selbsterfahrung zu tun.“

den Transfer in den Alltag und demnach die Wirksamkeit erlebnispädagogischer Maßnahmen zu sichern.

Die *Zielebenen* in der Erlebnispädagogik sind in der Literatur nicht eindeutig differenziert (vgl. Baig-Schneider, 2012, S. 19, S. 180; Reiner, 2007, S. 64ff.). Nach Schödlbauer (2000) können sie eingeteilt werden in persönlichkeitsbildende, soziale und interaktive, gesellschaftsintegrative sowie spirituelle oder religiöse Ziele. Unter den verschiedenen Zielsetzungen sind Begriffe zu finden, die dem *Kompetenzbereich* zugeordnet werden können: Bspw. Selbstvertrauen und Kritikfähigkeit unter persönlichkeitsbildenden Zielen; Kommunikation, Kooperation und Toleranz unter den sozialen und interaktiven Zielen (vgl. ebd., S. 72). Reiners (2007) unterteilt die Interaktionsspiele in drei Stufen, die auf die Entwicklung in unterschiedlichen Kompetenzbereichen abzielen.⁸² Im Englischen werden i.R. der *Experiential Education/Learning*, was der Erlebnispädagogik entspricht,⁸³ sog. *Hard Skills* und *Soft Skills* erworben, die Baig-Schneider (2012, S. 101) mit technischen (bspw. Klettertechnik) und sozialen Qualifikationen (bspw. Kommunikationsfähigkeit) übersetzt. Eine weitere Brücke, die zwischen *Experiential Education/ Learning* und SfE geschlagen werden kann, ist die Programmunterscheidung nach Priest und Gass (2005, S. 23) zwischen vier Stufen: Die erste Stufe *recreation* beinhaltet das Ziel, Gefühle zu verändern, *education* als Stufe zwei zielt auf die Veränderung von Denkweisen ab (*to change thinking*), Stufe drei *developmental* auf die Veränderung von Verhalten und die letzte Stufe *therapeutic* auf die Veränderung von *malbehaviour*. SfE-Ansätze sind auf der dritten Stufe anzusiedeln, wobei die zwei Stufen davor inkludiert sind.

⁸² Bei Stufe I liegt der Fokus auf der einzelnen Person, es werden Spiele eingesetzt, die das Selbstvertrauen stärken; Stufe II beinhaltet u.a. Kommunikationsspiele, die auf soziale Interaktionen abzielen; Stufe III beschäftigt sich bspw. mit Plan- oder Problemlösungsspielen, bei denen es darum geht, Erlerntes in eine „Welt außerhalb der Gruppe“ zu übertragen (vgl. Reiners, 2007, S. 39ff.).

⁸³ Im Englischen entsprechen die Bezeichnungen *Adventure-based Experiential Learning* oder *Experiential Education/ Learning* der Erlebnispädagogik, wobei die zweite Form definitorisch offener ist (Baig-Schneider, 2012).

2.3.3 Was ist der ‚deutsche‘ SfE-Ansatz? Präzisierung der sportbezogenen EZ-Maßnahmen

Auf Basis der aufgeführten definitorischen Ansätze und Modelle können zum Themenfeld SfE in Deutschland folgende allgemeingültige Aussagen getroffen werden:

Einhergehend mit gesellschaftlichen Veränderungen und dem damit verbundenen Wandel des Begriffsverständnisses von Sport (vgl. Blessing-Kapelke et al., 2007; Größing, 1993; Steinbach et al., 2008), hat sich auch der Sport im EZ-Kontext verändert. Im Sinne eines *weiten Sportverständnisses* und in Anlehnung an Wopp (1995) ist SfE heute der *freizeitsportlichen Ausprägung* des Sports mit dem Fokus sowohl auf einer breitensportlichen Ergebnisorientierung (Breitensport) und auf der soziokulturellen Erfahrungsorientierung, und damit dem *Erlebnissport*, zuzuordnen. Wurde der Sport früher ausschließlich aufgrund seiner ihm personen- und systembezogenen immanenten Funktionen im EZ-Kontext genutzt (vgl. Giebenhain, 1990, S. 189), wird er heute *im Sinne eines pädagogischen Auftrages (Bildung und Erziehung durch Sport) modifiziert* (bspw. durch Veränderung von Regeln und Aufgabenstellungen; Fokus auf Reflexionen von Erlebnissen während des Spiels) und dadurch *instrumentalisiert*, damit eine längerfristige Veränderung bzgl. Einstellungen und Verhalten i.R. der alltäglichen Lebenswelt gewährleistet wird. In einigen Partnerländern wurden nationale Sportlehrpläne für Schulen oder Lehrpläne für die Sportlehrerausbildung mitgestaltet oder ergänzt (Afghanistan, Indonesien, Namibia, Mosambik). Hier wurde ein Beitrag im Sinne der *Sportpädagogik* die *Erziehung im und durch Sport* durch die Gestaltung von ‚*ganzheitlichem Sportunterricht*‘ geleistet. Die Beiträge zu den Lehrplänen sind in ihrem Umfang und Inhalt sehr unterschiedlich und reichen von ganzheitlichen, kindgerechten und sportartunspezifischen Ansätzen (Afghanistan) bis hin zu fußballorientierten Beiträgen (Indonesien).

Des Weiteren möchte die dt. EZ mit den SfE-Maßnahmen *marginalisierte Zielgruppen* erreichen und ihnen den Zugang zu Sport ermöglichen (bspw. Mädchen und junge Frauen in Afghanistan, Kinder und Jugendliche mit Behinderungen in Uganda, Geflüchtete in Jordanien und Irak). Demnach orientiert sie sich an einer „Idee der universellen Inklusion der Bevölkerung in den Sport“ (Hartmann-Tews, 1996, S. 40), also dem *Sport-für-alle-Prinzip*.

Dadurch, dass die *Kompetenzorientierung* in deutschen Bildungsstandards eine große Rolle spielt, ist sie auch in der inhaltlichen Ausgestaltung der SfE-Maßnahmen zu finden: U.a. auf Grundlage des *Orientierungsrahmens für den Lernbereich Globale Entwicklung* (Schreiber & Siege, 2016) wurden sog. SfE-Kompetenzraster für Kinder und Jugendliche entwickelt, die als Standards genutzt werden sollen (vgl. S4D Resource Toolkit, o.D.d). Sie beinhalten eine Sammlung von persönlichen, sozialen und strategischen Kompetenzen in der Abstufung „Erkennen, Bewerten und Handeln“ (Schreiber & Siege, 2016, S. 90), die Teilnehmende durch SfE-Aktivitäten erlangen können.

Über die Sportpädagogik hinaus, gibt es speziell zur *Erlebnispädagogik* zahlreiche Überschneidungen: Neben den Schnittstellen bzgl. der Themen, die spezifische Zielgruppen einschließen (z.B. Inklusion, Migration, Gender) (vgl. Michl & Seidel, 2018), und des Erwerbs von Kompetenzen/sozialen Qualifikationen/Skills (vgl. Baig-Schneider, 2012, S. 101) ist vor allem die *Funktionslogik* dieselbe: Durch sportliche Aktivitäten und Spiele werden (1) *Ereignisse* erzeugt, (2) *Erlebnisse* reflektiert, die dadurch (3) zu *Erfahrungen* werden, die dann (4) in den Alltag transferiert werden sollen (vgl. Michl, 2009, S. 9f.). Dass gut gewählte *Reflexionsmethoden* hierbei entscheidend sind, um den Transfer in den Alltag und demnach die Wirksamkeit der Maßnahmen zu sichern, wird in der Erlebnispädagogik besonders betont (vgl. ebd.). Verschiedene Reflexionsmethoden werden allerdings bislang weder i.R. von SfE-Fortbildungen noch in den Handbüchern ausführlich behandelt.

Werden die sportlichen Aktivitäten der SfE-Maßnahmen genauer betrachtet, sind entweder *Spiele aus der Erlebnispädagogik* oder *klassische Mannschaftssportarten* mit modifizierten Trainingsinhalten zu finden. *Erlebnispädagogische (Interaktions-)Spiele* mit dem Ziel der Entwicklung von persönlichen und sozialen Kompetenzen sind Bestandteil der Handbücher in zahlreichen Partnerländern. Oft sind sie Ergänzungen in Manualen, die sich auf Sportarten fokussieren (bspw. Brasilien), teilweise aber auch Hauptbestandteil (bspw. Palästinensische Gebiete, Indonesien, Jordanien/*Manual Yalla Nela B* und *Region Westbalkan*) (vgl. S4D Resource Toolkit, o.D.c). Sie werden häufig als ‚*Kleine Spiele*‘

(*Small Games*) oder *Fun Games* bezeichnet und sind sportartenunspezifisch. Die folgenden Spiele sind Beispiele, die sowohl in der erlebnispädagogischen Literatur als auch in den SfE-Handbüchern zu finden sind:⁸⁴

- *Spinnennetz/Spider Web* zur Förderung von Kompetenzen, wie bspw. Kooperation, Kommunikation strategisches Denken und Problemlösung;
- *Giftfluss/Travessia/The floor is lava* zur Förderung von Kooperations-, Kommunikationsfähigkeit, Problemlösung und strategisches Denken;
- *Menschlicher oder Gordischer Knoten/Human oder Gordian Knot* zur Förderung von Kooperations-/Team-, Kommunikationsfähigkeit und Problemlösestrategien;
- *Wäscheklammern/Schwanzfangen/Grab the Goat's Tail/Tails Catch* zur Förderung von Selbstbewusstsein;
- *Blind/Roboter/Rely on me/Dinamica do Cego/Human Camera* zur Förderung von Kommunikationsfähigkeit, Verantwortungsbewusstsein und Vertrauen.

Dass neben diesen Interaktionsspielen, wie im internationalen EZ-Kontext auch (vgl. Schulenkorf et al., 2016), in erster Linie *klassische Mannschaftssportarten*, insbesondere Fußball, zum Zuge kommen (vgl. Schreiner, Mayer & Kastrup, 2020, S. 11), ist darauf zurückzuführen, dass dem Mannschaftssport an sich *positive soziale Lerneffekte* zugesprochen werden (bspw. Teamfähigkeit als Oberbegriff und Werte, wie Respekt, Toleranz und Fairplay) (vgl. Glöckner, 2010; Rolff, 2020). Auf dem Platz finden zu den Sportarten klassische Trainingsabläufe (Aufwärmen, Hauptteil, Cool Down) statt, in denen sportliche und nicht-sportliche Fähigkeiten (persönliche und soziale Kompetenzen) gefördert und themenspezifische Inhalte vermittelt werden.⁸⁵ Bestandteile sind meist sportartenspezifische Übungen und Spiele, die entweder *themenspezifisch modifiziert* werden, um Wissen zu vermitteln (bspw. Balldribbling zu Gender-Stereotypen) oder die

⁸⁴ An dieser Stelle wurden exemplarisch aus der Erlebnispädagogik die Spielsammlungen von Böhnke (2010) und Reiners (2007), exemplarisch Spiele aus den SfE-Manuale aus den Palästinensischen Gebieten (*Learn to Play – Play to Learn*; Steinbach & O'Rourke, 2017), Brasilien (*Treino Social 14+*; GIZ SV SfE, Juli 2016) und Jordanien (*Yalla Nela'B*; Muth & Wirsich, 2019) genutzt.

⁸⁵ Vgl. hierzu Trainingsabläufe und Übungsbeispiele in den Manualen aus Kolumbien (*Football/Deporte con Principios*), Jordanien (*Football/Basketball/Handball/Ultimate Frisbee builds Generations*), Äthiopien (*Sport2Work*), Namibia (*Football/Basketball/Volleyball/Netball4Life*), Kenia (*Drillbook Violence Prevention though Football*) im S4D Resource Toolkit (o.D.c).

kompetenzorientiert gestaltet sind (bspw. Dribbelaufgabe mit Partnerübung zu Kooperationsfähigkeit). Spielkonzepte der nationalen und auch internationalen Verbände finden sich wieder, wie bspw. die *Heidelberger Ballschule* (vgl. Roth & Kröger, 2015), *Kleine Spiele* des DFB oder *Kids' Athletics* (IAAF, 2006). Des Weiteren werden die Sportarten mit veränderten Regeln gespielt (bspw. *Crazy Soccer Tournament*, s. Steinbach & O'Rourke, 2017 oder *Football3*⁸⁶).

Bzgl. beider Schwerpunkte ist festzuhalten, dass die *Zusammenarbeit mit anderen deutschen SfE-Akteur*innen* großen Einfluss auf die Gestaltung der SfE-Methoden in den Partnerländern hat: Die Einbindung der (Dach-)Verbände (bspw. DOSB, DFB, DLV) beeinflusst die Wahl der Sportart; der Einbezug von Sportexpert*innen als Beratende bedingt die inhaltliche Ausrichtung und Gestaltung der SfE-Methoden. So waren bspw. DFB- und DOSB-Expert*innen maßgeblich an der Gestaltung in Brasilien, Jordanien, Kolumbien, Namibia, den Palästinensischen Gebieten, Kenia, der Türkei und Südafrika beteiligt, ein Berater von SPIN Sports Innovation mit erlebnispädagogischem Hintergrund in der Region West-Balkan, Indonesien und den Palästinensischen Gebieten (vgl. Schreiner et al., 2020).⁸⁷ Auch wenn Partnerorganisationen in den Ländern an der Entwicklung der SfE-Inhalte involviert waren (s. Impressen der Manuale), kann an dieser Stelle nicht nachvollzogen werden, in welchem Umfang eine partizipative Einbindung erfolgte und es stellt sich die Frage eines *weitergeführten ‚Kulturexports‘* (s. Kap. 2.1).

Deutlich wird, dass die *Bandbreite der deutschen SfE-Maßnahmen* sehr groß ist. Auch die einzelnen Komponenten in einem Land (bspw. Trainingsgestaltung, Eventformate) können sehr unterschiedlich sein. Würden die einzelnen Maßnahmen in den Partnerländern inklusive ihrer Angebote und spezifischer Methoden im Einzelnen analysiert, wäre eine weitere Differenzierung des SfE-Begriffs möglich. Die folgende Abbildung stellt einen ersten Versuch einer solchen Ausdifferenzierung dar, wobei der zweiendige Pfeil die

⁸⁶ *Football3* ist eine von *streetfootballworld* entwickelte Methode, bei der die Mitspieler*innen ohne Schiedsrichter*in spielen, eigene Regeln aufstellen und die dritte Halbzeit zur Reflexion und Diskussion des Spiels nutzen. Die Methode wird bspw. in Indonesien und in der Balkanregion genutzt (vgl. S4D Resource Toolkit, o.D.c).

⁸⁷ Basierend auf einer Dokumentenanalyse von deutschen Policy Dokumenten zu SfE von 2011 bis Mitte 2016 nehmen Schreiner et al. (2020) an, dass der DFB als hauptsächlicher Verbandspartner*in durch den Einsatz sog. ‚Sportexpert*innen‘ die operative Ebene, d.h. die konzeptionelle und damit inhaltliche Gestaltung der dt. SfE-Programme steuert.

Bandbreite symbolisiert und zeigt, dass die exemplarischen Maßnahmen des SV SfE der GIZ vom äußersten Ende von *Sport plus* bis zum äußersten Ende von *plus Sport* reichen.

Sport für Entwicklung (SfE)			
Sport plus Entwicklung		Entwicklung plus Sport	
Fokus: Sport (Turniere, Training) Primäre Akteur*innen: aus dem Sport		Fokus: Entwicklungsziele Primäre Akteur*innen: aus anderen (EZ-)Sektoren	
Bsp. Namibia: Fußballturniere des Namibischen Fußballverbands, bei denen Vorträge zu HIV/Aids-Prävention stattfinden und Teststände aufgestellt werden.	Bsp. Namibia: Trainingsaktivitäten des Projekts Galz&Goals mit einem Bildungsfokus und Kompetenzentwicklung hinsichtlich HIV/Aids-Prävention.	Bsp. Palästinensische Gebiete: Berufsbildungsworkshops für Jugendliche mit SfE-Trainingseinheiten zur Entwicklung von berufsbefähigenden Kompetenzen.	Bsp. Palästinensische Gebiete: Berufsbildungs-event mit einem SfE-Spielestand.

Abb. 5: Bandbreite deutscher SfE-Maßnahmen (eigene Darstellung in Anlehnung an DEZA, 2005)

Bei der genaueren Betrachtung der Maßnahmen fällt zudem auf, dass einige eher dem Bereich der *Sportförderung* zuzuschreiben wären, wie bspw. der Bau von Sportinfrastrukturen i.R. der BMZ-Initiative *Mehr Platz für Sport – 1000 Chancen für Afrika*. Viele haben auch eine *vordergründig Breitensportliche* Ergebnisorientierung, d.h. sie zielen nach dem *Sport-für-Alle*-Prinzip darauf ab, für marginalisierte Zielgruppen, die sonst eher schlechten oder keinen Zugang zum Sport haben, Sportangebote zu schaffen. Über die Zielgruppe (Mädchen, junge Frauen; Menschen mit Behinderung) wird bei diesen Projektansätzen eine Verknüpfung zu den Nachhaltigen Entwicklungszielen oder EZ-Schwerpunkten (Flucht, Inklusion, Gender etc.) hergestellt. Aber ist das dann *Förderung von Breitensport* oder *Sport für Entwicklung*? Letztendlich dreht sich alles um die Frage *„Wie viel Entwicklung steckt in den jeweiligen SfE-Maßnahmen?“*. Dass die Grenzen fließend sind, ist nicht neu und wird häufig betont (vgl. Coalter & Taylor, 2010, S. 15; DEZA, 2005, S. 19; Rockenfeller, 2009, S. 5;). Dies stellt allerdings keinen Grund dar, die Maßnahmen nicht genauer zu betrachten, um sie besser voneinander unterscheiden und definitorisch genauer verorten zu können. Schlussendlich hätte dies den Mehrwert, dass SfE-Ansätze und Methoden verfeinert und damit zielgerichteter zum Einsatz kommen könnten.

3. Rahmung: Theoretisch-konzeptionelle Untersuchungsperspektiven und methodische Vorgehensweisen

Folgendes Kapitel führt in unterschiedliche Kategorien ein, die den theoretisch-konzeptionellen Analyserahmen der Arbeit vorgeben. Grundlegende Evaluationsaspekte (Definitionen, theoretische Ansätze und Typen, Funktionen, Qualitätssicherung), Wirkungsmöglichkeiten und die Nachhaltigkeit von Projekten und Programmen als Evaluationsgegenstände sowie Akteur*innen im Evaluationsprozess mit dem Fokus auf der Rolle von Evaluierenden sind Bestandteile der Rahmung, die Kapitel 3.4 zusammenfasst. Das Kapitel schließt mit einem Einblick in die methodische Vorgehensweise der Arbeit.

3.1 Evaluation

3.1.1 Definitionen

Kromrey (2001, S. 105) spricht im Zusammenhang mit dem Begriff Evaluation von einem „*Diskussionslabyrinth*“: Evaluation ist vielfältig und komplex, nutzt unzählige Methoden und Verfahren, verfolgt eher forschende oder beratende Ansätze, betrachtet „ein kaum begrenzbares Spektrum von ‚Gegenständen‘“ und geht einer „unüberschaubaren Fülle von Fragestellungen“ nach. Der Begriff ist omnipräsent in verschiedensten Diskursen, „[...] im Alltag ebenso wie in der Politik, in der Methodologie empirischer Wissenschaft ebenso wie im (spezifischeren) Zusammenhang der Umfrageforschung“, wobei die „unterschiedliche(n) Kontexte [...] sich unterschiedliche Begriffsbedeutungen [schaffen]“ (ebd., S. 106).

Wie auch in dieser Arbeit, verwenden die meisten Autor*innen den Begriff der *Evaluationsforschung (evaluation research)* synonym mit *wissenschaftlicher Evaluation* (kurz: *Evaluation*) oder auch der *Evaluierung* (vgl. Caspari, 2004, S. 15; Döring & Bortz, 2016, S. 977, 979; Stockmann, 2000a, S. 11; Rossi et al., 2004, S. 6; Scriven, 2008, S. 139, 150; Vedung, 2000, S. 103), um zum Ausdruck zu bringen, „[...] dass es sich um *Bewertungen* handelt, die auf der Basis von Forschungsergebnissen *in einem wissenschaftlichen Pro-*

zess durch sozialwissenschaftlich qualifizierte Evaluationsfachleute vorgenommen werden“ (Döring & Bortz, 2016, S. 977).⁸⁸ Da der Begriff der Evaluation heute für jegliche Bewertungen, die wissenschaftlichen Standards allerdings nicht gerecht werden, recht „inflationär“ verwendet wird (vgl. Beywl & Widmer, 2009, S. 15; Kromrey, 2001, S. 105), ist die *Forschungsbetonung* wichtig, sie wird zugleich jedoch häufig hinterfragt oder kritisiert (vgl. Kromrey, 2003; Lüders, 2006). Im Englischen wird auch häufig der Begriff *programme evaluation* genutzt, was den Hauptgegenstand von Evaluationen beinhaltet (vgl. Mouton, 2014; Rossi et al., 2004; Scriven, 2008, S. 285).⁸⁹

Im Gegensatz zur Grundlagenforschung und Interventionsforschung, die sich mit der Entwicklung und Überprüfung wissenschaftlicher Theorien befassen, operiert die Evaluationsforschung meist stärker *theorieanwendend* als -entwickelnd. Bereits bestehende Theorien werden zur Bewertung von Evaluationsgegenständen genutzt und dementsprechend ausgewählt. Demnach ist die Evaluationsforschung eine *anwendungsorientierte Forschung*, die immer mit einem *Nutzen* verbunden ist, d.h. es soll „Handlungswissen für die Praxis“ geschaffen werden (ebd., S. 113). Nach Vedung (2000, S. 124) liegt hier die *Hauptunterscheidung zwischen Grundlagen- und Evaluationsforschung*: „[...] the basic difference between evaluation research and fundamental research is that the former is *intended for use*.“⁹⁰ Sie ist damit auch *weniger frei* als bspw. die Grundlagenforschung, die zweckfrei nach Erkenntnissen streben kann. Evaluationsforschung richtet sich meist an den Fragestellungen, Absichten, Zeit- und Budgetvorgaben eines Auftraggebenden aus und impliziert immer eine Bewertung anhand bestimmter Kriterien als Teil des Forschungsauftrages (Döring & Bortz, 2016, S. 977; vgl. Stockmann, 2000a, S. 12; 2007b, S. 28).⁹¹

Nach Dahler-Larsen (2017, S. 22) umfassen *Evaluationsdefinitionen* typischerweise vier Komponenten: (1) die systematische Untersuchung; (2) Kriterien; (3) Evaluationsgegenstände (bspw. Policies, Programme, Projekte, Interventionen) und (4) einen intendierten

⁸⁸ Wollmann (2017, S. 2) nimmt bspw. eine Unterscheidung zwischen den Begriffen vor. In dieser Arbeit werden die Begriffe synonym verwendet und vornehmlich der Begriff der Evaluation genutzt.

⁸⁹ S. Kapitel 3.2.

⁹⁰ Hervorhebungen durch die Autorin.

⁹¹ Mouton (2014, S. 165) führt die Hauptunterschiede zwischen *evaluation* und *research* in einer Tabelle auf.

Zweck oder Nutzen (bspw. Rechenschaftslegung, Optimierung, Lernen, Entwicklung). Eine umfangreiche Definition nach Döring und Bortz (2016, S. 979) greift alle vier Komponenten auf und geht auf weitere zentrale Aspekte ein:⁹²

„Die Evaluationsforschung (‘evaluation research’) bzw. wissenschaftliche Evaluation (‘evaluation’) nutzt sozialwissenschaftliche Methoden⁹³, um einen Evaluationsgegenstand (z.B. ein Produkt oder eine Maßnahme) unter Berücksichtigung der relevanten Anspruchsgruppen (z.B. [...] Evaluationsauftraggeber) anhand bestimmter Evaluationskriterien [...] zu bewerten. Die durch den Prozess der Evaluationsforschung im Ergebnis erlangte Bewertung soll in der Praxis unterschiedliche Evaluationsfunktionen erfüllen (z.B. Legitimations- oder Optimierungsfunktion), was eine aktive Evaluationsnutzung, d.h. ein praktisches Aufgreifen der [...] Evaluationsergebnisse verlangt [...]. Die Evaluationsforschung ist den üblichen Prinzipien der Wissenschaftlichkeit sowie der Wissenschafts- und Forschungsethik verpflichtet, die für die besonderen Bedingungen der Evaluationsforschung zusätzlich in verbindliche Evaluationsstandards als Gütestandards für Evaluationsstudien gefasst werden. Die Evaluationsforschung stellt über sozialwissenschaftliche Methodenkompetenz hinaus besondere Anforderungen an die Qualifikation der Evaluierenden.“

Dass Evaluation *Teil der angewandten Sozialforschung* ist und hierfür die „gesamte Bandbreite der sozialwissenschaftlichen Theorien, Konzepte und Forschungsmethoden“ (Stockmann, 2007a, S. 18) genutzt wird, wird in der Literatur und in Evaluationsstandards betont (vgl. Caspari, 2004, S. 94; DeGEval, 2016, S. 46; Döring & Bortz, 2016, S. 976; Rossi et al., 2004, S. 16f.; Stockmann, 2007b, S. 28). Auch die *Qualifizierung von Evaluierenden* ist ein wichtiger Aspekt, der u.a. *professionelle Evaluationen* von *Laienevaluationen* oder *Alltagsevaluationen* unterscheidet, „bei denen irgendetwas von irgendjemanden nach irgendwelchen Kriterien in irgendeiner Weise bewertet wird“ und Evaluationen zu einer *Gefahrenquelle* werden können, indem Sachverhalte bspw. einseitig oder falsch dargestellt werden (ebd., S. 27; vgl. auch Döring & Bortz, 2016, S. 992ff.; Kromrey, 2001, S. 107). Weitere Aspekte der Definition (Evaluationsgegenstände, Anspruchsgruppen, Evaluationsfunktionen und -nutzung, Evaluationsstandards) sind Gegenstände weiterer Teilkapitel.

⁹² Weitere Definitionen verschiedener Autor*innen s. Auflistungen bei Caspari (2004, S. 17), Döring und Bortz (2016, S.979) sowie Stockmann (2000a, S. 12f.; 2007, S. 26f.)

⁹³ Unter *Methode* wird die „systematische, d.h. nach festgelegten Regeln ablaufende Vorgehensweise der Sozialforschung vor allem zur Datengewinnung und -auswertung“ verstanden (Oberndörfer et al, 2010, S. 240).

In Zusammenhang mit Evaluation werden weitere Begriffe wie *Appraisal*, *Assessment*, *Auditing* und *(Financial) Controlling*⁹⁴ genutzt oder, wie mit dem Begriff *Monitoring*, kombiniert (*Monitoring und Evaluation*, kurz: *M&E*). Auch wenn sich die mit den Begrifflichkeiten verbundenen Aktivitäten von jenen der Evaluation unterscheiden, sind die Trennlinien unscharf (Berlage & Stokke, 1992; Scriven, 2008). Anders als Evaluationen, die zu einem bestimmten Zeitpunkt durchgeführt werden, ist *Monitoring* für Stockmann (2006) eine deskriptive Daueraufgabe, was die kontinuierliche Sammlung von Informationen/Daten, bspw. i.R. eines kompletten Programmverlaufs, beinhaltet. Evaluationen lassen sich seiner Meinung nach allerdings zu einem „kontinuierlichen Monitoring ausbauen“ (ebd., S. 75). Kromrey (2001, S. 123) bezeichnet diese „systematische Buchführung“ als eine Hauptaufgabe von Evaluation.

Die *wissenschaftliche Begleitung* bzw. *Begleitforschung* und *Evaluation* werden oft häufig synonym (vgl. Kind & Wessels, 2008, S. 347) oder parallel verwendet (vgl. Schemme, 2003), meinen nach Kromrey (2007, S. 113) jedoch „fast das Gleiche, und doch etwas Anderes“, obgleich bei ihm eindeutige Definitionen, die Gemeinsamkeiten und Unterschiede aufzeigen, nicht zu finden sind. Kind und Wessels (2008, S. 348) bezeichnen die Begleitforschung als „[...] ein Bündel an Einzelmaßnahmen, die zusammen die Programm- und Projektumsetzung begleiten und unterstützen soll“. Sie beinhaltet sehr unterschiedliche Aufgaben und hat damit verbundene, komplexe Erwartungen zu erfüllen (ebd.). Nach Schemme (2003, S. 31) hat eine wissenschaftliche Begleitung vier Funktionen: (1) die theoretische und methodische Fundierung in der Planungsphase, (2) prozessorientierte Begleitung und Optimierung der Gestaltung und Implementation neuer Konzepte (Dialog, Beratung, Moderation, Mediation, formative Evaluation); (3) Dokumentation (Beschreibung, Analyse und Bewertung von Verlauf und Ergebnissen), Öffentlichkeitsarbeit (Außenwirkung, Verbreitung, Transfer), Nachweis der Relevanz und

⁹⁴ Berlage und Stokke (1992, S. 6) definieren *Appraisal* als ex-ante Evaluation oder als eine finale Untersuchung einer Maßnahme bzgl. ihrer Relevanz. Sie entscheidet darüber, ob eine Maßnahme letztendlich durchgeführt wird oder nicht. Mit einem meist von externen Evaluierenden durchgeführten *Auditing* soll überprüft werden, ob die Maßnahmen, Prozesse und Verfahren des Gebers den im Voraus festgelegten Normen und Kriterien entsprechen (ebd.). *Controlling* ist ein betriebswirtschaftlicher Ansatz, der dem Management dabei hilft, „[...] das Gesamtsystem ergebniszielorientiert an Umweltveränderungen anzupassen“ (Stockmann, 2006, S. 76) und weist viele Ähnlichkeiten zum Monitoring auf (vgl. ebd., S. 76ff.). S. Definitionen auch bei Scriven (2008, S. 58, 60, 62).

Wirksamkeit des Vorhabens, summative Evaluation; (4) kooperative Gewinnung neuer Erkenntnisse. Demnach sind Evaluationen Gegenstand oder ein integrierter Teil wissenschaftlicher Begleitung/Begleitforschung, sie umfassen allerdings zusätzlich weitere Aufgaben und Arbeitsbereiche (Beratung, Vernetzung etc.).

3.1.2 *Evaluationsansätze und -klassifikation*

Evaluationsansätze, die sich im Laufe der Zeit entwickelt haben, „[...] orientieren sich entweder mehr an wissenschaftlichen Standards oder stärker an den Anforderungen der Auftraggeber oder den Bedürfnissen der Zielgruppen“ (Stockmann, 2007b, S. 29). D.h. auch in den Evaluationsansätzen zeichnet sich die besondere Dualität bzw. das Spannungsverhältnis zwischen Wissenschaftlichkeit und Nützlichkeit von Evaluation ab.

Es existieren *unzählige Konzepte, Modelle, Theorien und Ansätze* von Evaluation in der Theorie und Praxis, die u.a. Handlungsanweisungen bzgl. des Evaluationsdesigns und der -durchführung beinhalten (Shadish, Cook & Leviton, 1991). Genauso viele Versuche wurden unternommen, alternative konkurrierende Evaluationskonzepte, -modelle, -theorien und -paradigmen auf der Grundlage ihrer Ähnlichkeiten und Unterschiede zu klassifizieren. Basierend auf bekannten Klassifizierungssystemen,⁹⁵ präsentiert Rabie (2014) eine weitere umfassende Einteilung, die Begrenzungen der genutzten Systeme kompensieren soll. Sie wird für diese Arbeit aufgrund ihrer umfassenden und klaren Struktur genutzt und z.T. ergänzt. Hier werden drei Hauptkategorien zur Klassifizierung genutzt, die auf dem zugrundeliegenden konzeptionellen, theoretischen oder praktischen Ansatz oder der treibenden Kraft hinter der Evaluation beruhen, die die Grundlage für die Evaluation darstellt:⁹⁶

⁹⁵ Darunter der *Evaluation Theory Tree* nach Alkin und Christie (2009), der bekannte Evaluationstheoretiker*innen in die Kernkomponenten *methods*, *use* und *valuing* unterteilt und einem System nach Rossi et al. (2004), die Evaluation anhand des Programmzyklus einordnen. Weitere Übersichten geben Astbury (2017) und Stockmann (2007). Details zur Entstehung und Rolle von Evaluationstheorien s. Donaldson und Lipsey (2007).

⁹⁶ Die wichtigsten Ansätze können nicht ausführlich beschrieben, sondern lediglich umrissen werden, wobei nicht immer die Ursprungsliteratur genutzt wird. Ausführliche Beschreibungen und Ursprünge s. Rabie (2014). Bei der für die Arbeit relevantesten Ansätze (z.B. Logical Framework), wird weitere Literatur ergänzt.

(1) Ansätze basierend auf dem (Anwendungs-)Bereich oder Umfang (*scope*)

Evaluationen können sehr breit bzw. umfassend sein, nur Teilbereiche von Maßnahmen oder einen einzelnen Gegenstand in den Fokus rücken. In diesem Zusammenhang können sie einen eher *bilanzierend-summativen* Charakter haben, der bspw. die gesamte Maßnahme im Blick hat oder einen eher *formativ-gestaltenden* bzgl. bestimmter Gegenstände oder Teilbereiche (Döring und Bortz, 2016).⁹⁷ Darüber hinaus können Ergebnisse von Evaluationsstudien auch zusammengefasst werden, gerade vor dem Hintergrund, dass die Anzahl von Evaluationen stetig steigt (Caspari, 2015).

Der Begriff *Review* steht in diesem Zusammenhang für einen Oberbegriff, der „[jedwede] Art einer zusammenfassenden Darstellung verschiedener Publikationen, Studien oder Evaluationen zu einem Themenbereich [meint]“ (Caspari, 2015, S. 149)⁹⁸. Der Begriff *Evaluationssynthese* steht „[...] für die Querschnittsauswertung und Synthese von verschiedenen Evaluationsstudien sowie weiteren Quellen [...]“ (ebd.) (z.B. andere Forschungsstudien, Berichte) zu einem gemeinsamen Gegenstand/Evaluandum, um generelle Schlussfolgerungen und (maßnahmen-)übergreifende Erkenntnisse zu erhalten (z.B. über Reichweite, Effektivität, Wirksamkeit). *Qualitative* Formen werden als *Meta-Synthesen* (Booth, 2016) oder *narrative Reviews* bezeichnet. *Quantitative* Evaluationssynthesen sind sog. *Meta-Analysen*. Diese Analyse von Analysen ist ein rein statistisches Syntheseverfahren, bei dem homogene quantitative Primärdaten zu Metadaten aggregiert werden. Eine *Systematic Review* ist ein „hochsystematisches Analyse- und Synthese-Verfahren, bei dem vornehmlich quantitative Befunde vergleichbarer empirischer Einzelstudien zusammengefasst und deren Variabilität untersucht werden“ (Caspari, 2015, S. 150). Vorab wird eine explizit definierte Methodik gewählt, d.h. eine eindeutige Suchstrategie bzgl. relevanter Studien, klar definierte Auswahlkriterien, „[...] eine systematische Kodierung und Analyse der berücksichtigten Studien sowie eine Metaanalyse, sofern möglich“. Im Unterschied zur Evaluationssynthese werden bei einer Systematic

⁹⁷ Die zentrale Unterscheidung zwischen *formativer und summativer Evaluation* geht auf den bekannten Evaluationsforscher *Michael Scriven* zurück. S. ausführliche Definitionen bei Scriven (2008, S. 168f., 340).

⁹⁸ Caspari (2015) stützt sich bei ihrer Einordnung u.a. auf Scriven (2008) und weitere internationale Autor*innen.

Review vornehmlich quantitative Verfahren genutzt, „[...] um anspruchsvolle Primärdaten einer Sekundärdatenanalyse zu unterziehen“ (bspw. von randomisierten kontrollierten Studien, RCTs)⁹⁹. Eine *Meta-Evaluation* evaluiert Evaluationen. Dieses Review-Verfahren analysiert anhand definierter Bewertungskriterien (bspw. Checklisten die auf Evaluationsstandards basieren), „[...] inwieweit eine oder mehrere Evaluationen Standards erfüllen und [bewertet] die Qualität einer oder mehrere Evaluationen [...]“ (ebd.; Scriven, 2008).¹⁰⁰

Zur Evaluation von *Teilbereichen von Maßnahmen* zählen u.a. Input- oder Machbarkeitsstudien, Prozessevaluationen, Output, Outcome und Impact bzw. *Wirkungsevaluationen* (Rabie, 2014). Die Evaluationen orientieren sich an verschiedenen Phasen einer Maßnahme, wobei sie andere Analyseperspektiven einnehmen und unterschiedliche Erkenntnisinteressen verfolgen, wobei auch von *ex-ante*, *on-going*, *finalen* und *ex-post* Evaluationen gesprochen wird (vgl. Caspari, 2004, S. 28ff.; Döring & Bortz, 2016, S. 991; Rossi et al., 2004, S. 40; vgl. Stockmann, 2007b, S. 33ff.; 2006, S. 69ff.).

Ex-ante Evaluationen (auch *preformative Evaluationen*) werden in der Konzeptionierungs- oder Planungsphase eines Programms durchgeführt und haben zum Ziel, die gewonnenen Erkenntnisse für die weitere Entwicklung und Gestaltung der Maßnahme zu nutzen. Das Erkenntnisinteresse liegt hierbei bspw. auf der Analyse von Bedarfen und der Untersuchung der Ausgangslage. Hierzu gehören auch *Grundlagenuntersuchungen* bzw. sog. *Baseline Studies*. *On-going Evaluationen* (auch *Mid-term-* oder *formative Evaluationen*) untersuchen den Programmverlauf in der Implementationsphase. Hierbei werden kontinuierlich Informationen gesammelt (*Monitoring*), systematisiert und bewertet – mit dem Ziel, das Programmmanagement bei der Steuerung zu unterstützen, bspw. werden Ergebnisse zur Entscheidungsfindung oder für Kurskorrekturen genutzt.

⁹⁹ S. Erläuterung unter (3) *Evaluationsdesign und Methoden*.

¹⁰⁰ Weitere Formen von umfassenderen Evaluationen, die *gesamte Maßnahmen* im Blick haben, sind bspw. die *Policy Evaluation*, die öffentliche Policies im Hinblick auf ihre Effektivität bewertet oder verschiedene Policies vergleicht (Rabie, 2014), oder die *systemische Evaluation* (s. Margo & van den Berg, 2019), bei der der Fokus auf einem gesamten System oder Subsystem liegt, d.h. dass das Verständnis über das System, das eine Maßnahme umgibt, von Evaluationsinteresse ist (Vaessen, 2019). Des Weiteren nennt Rabie (2014, S. 122ff.) *Community Evaluation*, *Assessments zur Evaluierbarkeit* und *Programme Monitoring und Evaluation*. Letzteres wird seitens der Autorin nicht als ein separater Ansatz verstanden und deswegen an dieser Stelle nicht aufgeführt.

Finale oder sog. *Schlussevaluationen* werden direkt am Ende oder ‚als letzte Aktion‘ der Implementationsphase einer Maßnahme durchgeführt. Sie überprüfen bilanzierend/summativ, inwieweit eine Maßnahme entsprechend der Planung umgesetzt und Ziele erreicht wurden.¹⁰¹ *Ex-post Evaluationen* werden nach Abschluss einer Maßnahme oder einige Zeit nach Förderende durchgeführt. Sie haben die Aufgabe, die *Gesamtwirkungen*, die durch das Vorhaben ausgelöst wurden, zu erfassen, zu bewerten und Zusammenhänge aufzudecken sowie dessen *Nachhaltigkeit* zu untersuchen. Im besten Falle werden die Erkenntnisse für künftige Vorhaben genutzt. Demnach haben ex-post Evaluationen nicht nur einen summativen, sondern auch einen preformativen oder formativen Charakter.¹⁰²

(2) Ansätze basierend auf einer expliziten Philosophie oder Theorie

Hier wird unterschieden zwischen *theoriegeleiteten Ansätzen*, die auf der Basis fundierter Sachkenntnisse zu den zu untersuchenden Phänomenen das Ziel verfolgen, das Wissen über den Evaluationsgegenstand zu erweitern, und *partizipativen Ansätzen*, die einen eher sozialwissenschaftlichen Ansatz zur Evaluation bevorzugen, mit dem allgemeinen Ziel, Beteiligte zu befähigen (*Empowerment*) und ein gemeinsames Verständnis zu schaffen. Zu den *theoriegeleiteten (theory-driven) Evaluationsansätzen* zählt u.a. der Ansatz des *Logischen Modells*.¹⁰³ Die Begriffe *Logical Framework*, *Logframe*, *Logisches Modell (Logic Model)*, *Wirkungslogik*, *Theory of Change*, *Wirkungsmodell (results model)*, *Wirkungskette (results chain)* und *Programmtheorie (programme theory)* stehen hierbei für dasselbe: Sie haben die Aufgabe, die Funktionsweise eines Projekts oder Programms und den kausalen Zusammenhang zwischen *Inputs*, *Outputs*, *Outcomes* und *Impacts* zu veranschaulichen und zu erklären (vgl. Oberndörfer et al., 2010, S. 21; OECD, 2009; Kurz

¹⁰¹ Fälschlicherweise werden solche Evaluationen in der EZ oft als *ex-post* bezeichnet (bspw. MacArthur, 1995, S. 3f., 20; Cracknell, 2000, S. 74). Als *ex-post* Evaluationen werden in Anlehnung an Caspari (2004, S. 30) jene Untersuchungen verstanden, die mit einem zeitlichen Abstand nach Ende der Förderung vorgenommen werden.

¹⁰² Auch wenn *ex-post* Evaluationen nach Rabie (2014) Einteilung zu einem Ansatz zur Evaluation von gesamten Maßnahme gezählt wird, wird er an dieser Stelle in Zusammenhang mit den *ex-ante*, *on-going* und *finalen* Evaluationen genannt, da sie aus analytischer Perspektive sich auf ausgewählte Aspekte fokussieren (Wirkungen, Nachhaltigkeit) und da die Gesamtstruktur dieser Einteilung in ihrer Gänze dargestellt werden soll.

¹⁰³ Nach Rabie (2014, S. 129ff.) zählt das *Logical Framework* zu den *klärenden (clarificatory)* bzw. *konzeptionellen (conceptual)* Evaluationsansätzen als Unterkategorie der *theoriegeleiteten Ansätze*.

& Kubek, 2018, S. 34; Scriven, 2008, S. 286);¹⁰⁴ Als *Inputs* werden finanzielle, personelle oder materielle Ressourcen verstanden, die bei der Durchführung von Maßnahmen eingesetzt werden. *Outputs* bezeichnen alle Leistungen, d.h. die Angebote und Produkte einer Maßnahme (bspw. Anzahl an betreuten Kindern, gebauter Sportplätze etc.). *Outcomes* sind kurz- und mittelfristige Wirkungen auf Ebene der Zielgruppe einer Maßnahme, resultierend aus den *Outputs*. Mit *Impacts* sind *längerfristige Wirkungen bzw. Veränderungen auf der Ebene der Gesellschaft* gemeint. Sie beinhalten „positive und negative, primäre und sekundäre Langzeiteffekte, die direkt oder indirekt, beabsichtigt oder unbeabsichtigt“ durch eine Maßnahme verursacht werden (vgl. OECD, 2009, S. 31; Stockmann, 2006, S. 101f.; Kurz & Kubek, 2018, S. 35ff.). Abbildung 6 stellt ein lineares *Logisches Modell* dar, das nach Kurz und Kubek (2018) die am weitesten verbreitete Version einer Wirkungslogik ist. *Indikatoren* als Maß- oder Messeinheiten werden für jede Ebene formuliert. Sie können in ihrer Ausprägung *zählbar* (bspw. ‚Anzahl der Jugendlichen, die am Projekt teilnehmen‘ als *quantitativer Output-Indikator*) oder *beschreibbar* (bspw. ‚Jugendliche haben eine klare berufliche Perspektive‘ als *qualitativer Outcome-Indikator*) sein (ebd.).¹⁰⁵

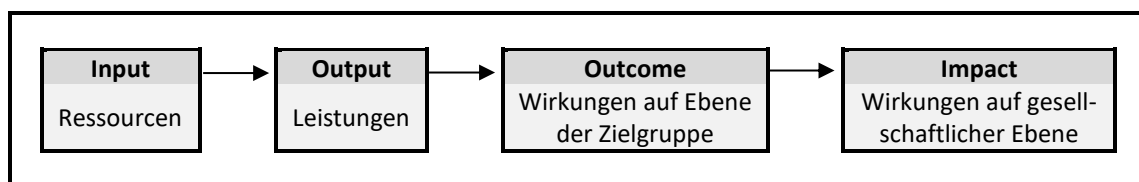


Abb. 6: Logisches Modell/Logical Framework (vgl. Kurz & Kubek, 2018, S. 35)

Zu den *partizipativen (participatory) Evaluationsansätzen* gehören u.a. die *critical theory evaluation*, die darauf abzielt, den Wert einer Sache zu bestimmen, indem falsche oder kulturell begründete Sichtweisen durch einen systematischen Untersuchungsprozess aufgedeckt werden. Hierzu gehört die sog. *feministische Evaluation*, die auf die Förderung von sozialer Gerechtigkeit für Frauen abzielt; die *naturalistische/konstruktivistische Evaluation*, die versucht den Evaluationsprozess in das Leben der beteiligten Personen

¹⁰⁴ S. hierzu auch Hemmer (2007), ausführlich zu Prozessen bzw. Phasen Binnendijk (2007) und zu Unterschieden und Gemeinsamkeiten von Logic Models, Logical Frameworks und RBM Cummings (1997).

¹⁰⁵ Weitere theoriegeleitete Ansätze sind die *realistische (realist/realistic) Evaluation*, die nach Pawson und Tilley (1997) die Bedeutung des *Kontexts* betont oder die *entwickelnde (developmental) Evaluation*, die soziale Innovationen und adaptives Management, also die Anpassungsfähigkeit an eine komplexe Umwelt unterstützt (Patton, 2011). Des Weiteren nennt Rabie (2014) die *zielfreie (goal-free)* und die *erhellende (illuminative) Evaluation*.

einzubinden; die *Empowerment Evaluation*, die auch die Befähigung von Beteiligten beinhaltet, damit diese selbst Evaluationen durchführen können und die *nutzungs- bzw. nutzerorientierte (utilization-focused) Evaluation*, die Nutzer und die Nutzung von Evaluation in den Fokus rückt (Patton, 2008).¹⁰⁶

(3) Evaluationsdesign und -methoden

Das *Evaluationsdesign* beschreibt die Art und Weise sowie den Prozess der Datenerhebung und -auswertung und damit „die gesamte Gestaltung des Vorgehens zur Beantwortung einer (wissenschaftlichen) Fragestellung. Dazu gehört die Auswahl von Verfahren und Instrumenten [*Methoden*] und der Zeitplan unter den gegebenen Restriktionen (Geld, Zeit, Personal, Raum)“ (Oberndörfer et al., 2010, S. 240). Es wird unterschieden zwischen quantitativen, qualitativen und gemischtmethodischen, sog. Mixed Methods Designs:

Quantitative Designs umfassen klassische *experimentelle* und *quasi-experimentelle* Designs. Bei experimentellen Designs mit Kontrollgruppen, die einen ‚mit-ohne‘-Vergleich implizieren, werden Veränderungen bei einer vergleichbaren Kontrollgruppe, die nicht Teil der Maßnahme sind, mit jenen der Zielgruppe verglichen. Bei sog. *Randomized Controlled Trials (RCT)* werden Individuen vor der Implementierung einer Maßnahme nach dem Zufallsprinzip (Randomisierung) in eine Ziel- und eine Kontrollgruppe zugeordnet. *Quasi-experimentellen Designs* hingegen orientieren sich an der Experimentallogik, wobei eine Vergleichsgruppe (re-)konstruiert wird.¹⁰⁷ Im Unterschied zum experimentellen Design erfolgt die Bestimmung der Vergleichsgruppen im Nachhinein, d.h. entweder im Laufe der Maßnahme oder gar erst nach Abschluss der Maßnahme. Im Bereich von Wirkungsevaluationen beantworten diese Designs insbesondere die Frage, *ob* und in welchem Umfang Wirkungen eingetreten sind (Caspari & Barbu, 2008; Mark & Henry, 2007;

¹⁰⁶ Des Weiteren nennt Rabie (2014, S. 136ff.) die *wertschätzende Evaluation*, die sich auf die Stärken eines Evaluationsgegenstandes fokussiert, die *responsive Evaluation*, die persönliche Erfahrungen in den Mittelpunkt stellt und die *demokratische Evaluation*, die sich auf Konzepte der Demokratie stützt, um unter Berücksichtigung aller relevanten Interessen, Werte und Perspektiven zu vertretbaren bewertenden Schlussfolgerungen zu gelangen.

¹⁰⁷ Zur Konstruktion von Vergleichsgruppen gibt es unterschiedliche Möglichkeiten sog. *Matching-Verfahren*, s. hierzu Caspari und Barbu (2008, S. 13f.)

Müller, 2017; Neubert, 2010; Oberndörfer et al., 2010; Reade, 2008; Roxin, 2016; Scriven, 2008). *Qualitative* (nicht-experimentelle) Designs gehen über die Evaluation von Programmergebnissen und -wirkungen hinaus. Sie beantworten insbesondere Fragen nach dem *warum* und *wie* und konzentrieren sich auf den konstruierten Charakter einer Maßnahme, d.h. deren Einbettung in und Beeinflussung durch sie umgebende Kontexte oder Systeme, sowie auf die Wichtigkeit von Umsetzungsprozessen (Caspari & Barbu, 2008; Rabie, 2014). Beispiele für solche Designs sind sog. *rapid appraisals*, bei denen qualitative Methoden (Fokusgruppeninterviews, Beobachtungen, Experteninterviews) genutzt werden, um lokale Gegebenheiten und Bewohner*innen zu untersuchen, *Umfragen (surveys)* und *Fallstudien*, bei denen bspw. Interviewtechniken eingesetzt werden oder der Ansatz der *participatory action research (PAR)*, der Evaluationsprozesse mit der Aufklärung von „less powerfull“ Interessengruppen und anschließenden Maßnahmen auf der Grundlage der Forschungsergebnisse beinhaltet (ebd., S. 148).¹⁰⁸ *Mixed-Methods* Designs kombinieren quantitative und qualitative Verfahrensweisen, um ein Phänomen aus unterschiedlichen Perspektiven zu beleuchten und die jeweiligen ‚Schwächen‘ von Methoden auszugleichen. Durch *Triangulation* wird hierbei die Stimmigkeit der mit verschiedenen Instrumenten gewonnenen Erkenntnissen geprüft, um Ursachen zu ermitteln, die die Ergebnisse beeinflussen (ebd.; Baur, Kelle & Kuckartz, 2017; Creswell & Creswell, 2018).¹⁰⁹

¹⁰⁸ PAR und weitere partizipative Ansätze gehen zurück auf den *Participatory Rural Appraisal (PRA)*, die „Mutter aller partizipativen Ansätze“ (Bliss, Neubert & Heint, 2007, S. 191). Der Ansatz soll lokale Menschen dazu befähigen, ihr Wissen bzgl. ihrer Lebensumstände zu teilen, zu erweitern und zu analysieren, und sie dadurch handlungsfähiger zu machen (vgl. Chambers, 1994, S. 953). In *‘The Origins and Practice of Participatory Rural Appraisal’* beschreibt Chambers Ursprünge, Prinzipien, Ansätze, Methoden und Anwendungen von *PRA* (ebd.).

¹⁰⁹ Die Autor*innen beschreiben unterschiedliche Mixed-Methods Designs, wie dem *konvergenten*, dem *erklärenden sequentiellen* und dem *explorativen sequentiellen* Design, bei denen quantitative und qualitative Methoden in unterschiedlicher Reihenfolge zum Einsatz kommen (vgl. Creswell & Creswell, 2018, S. 300).

3.1.3 Zur Funktion bzw. zum Nutzen von Evaluation

Evaluation dient „[...] stets einem von außen an sie herangetragenen Zweck“, d.h. sie wird in den „Dienst der Gesellschaft“ gestellt und hat verschiedene *Funktionen* zu erfüllen, die, wie auch die Wahl der Evaluationsthemen, stark vom jeweiligen Interesse der Auftraggebenden bestimmt werden (Stockmann & Meyer, 2017, S. 156; Stockmann, 2017). Die Funktion von Evaluation steht in engem Zusammenhang mit dem *Evaluationszweck (purpose for evaluation)* (Stern, 2005 zitiert in Vestman & Conner, 2007, S. 230), dem *Nutzen bzw. der Nutzung von Evaluation (use of evaluation)* (vgl. Vedung, 2015; Weiss, 1998).¹¹⁰ „Welchen Zweck bzw. Nutzen Evaluationen verfolgen oder welche Funktion sie haben, ebenso wie die zugrundeliegenden Konzepte, ist in der Literatur bisher noch nicht einheitlich dargestellt worden“ (Caspari, 2004, S. 18). Je nach Kontext und Perspektive ändern sich Schwerpunkte und Bezeichnungen. Die Funktions- und Nutzungszuschreibungen unterscheiden sich in ihrer Benennung und Anzahl, weisen allerdings zahlreiche Überschneidungen auf. Im Folgenden erfolgt der Versuch einer Zuordnung:

Die *Erkenntnisfunktion* wird häufig als zentrale Funktion von wissenschaftlichen Evaluationen gesehen (vgl. Stockmann, 2007b, S. 36). Evaluationen sollen in erster Linie wissenschaftliche, d.h. auch möglichst objektive Erkenntnisse über die Eigenschaften und Wirkungen von Evaluationsgegenständen eruieren. Sie zielen hierbei auf eine Informationsvermittlung ab, d.h. das Wissen unterschiedlicher Nutzergruppen soll bzgl. des Evaluationsgegenstandes erweitert werden (Döring & Bortz, 2016, S. 987; Rossi et al., 2004, S. 36f.¹¹¹). Spiel und Schober (2015) betonen in diesem Zusammenhang die Gesellschaft als zentrale Nutzergruppe und sprechen von *gesellschaftlicher Aufklärung*. Die Funktion deckt sich mit dem *konzeptuellen/aufklärenden* Verwendungstyp von Evaluationen (vgl. Müller, 2015, S. 29), in englischsprachiger Literatur als *conceptual* oder *enlightenment use* bezeichnet. Hier wird betont, dass dieser „weaker type of utilization“ (Vedung, 2015, S. 194) dazu beiträgt, dass Nutzergruppen „new ideas and insights“ (Weiss, 1998, S. 24) bzgl. des Evaluationsgegenstandes erlangen, was das Denken von Nutzern beeinflussen

¹¹⁰ Die Nutzung und der Nutzen kann auch bei Evaluationsansätzen in den Vordergrund gerückt werden, s. *utilization-focused evaluation* nach Patton (2008).

¹¹¹ Rossi et al. (2004, S. 34) sprechen von *knowledge generation*.

kann, ohne allerdings weitere Entscheidungen zu treffen oder in Aktion zu treten (vgl. ebd.; Vedung, 2015, S. 194; Rossi et al., 2004, S. 432ff.).

Die *instrumentelle Nutzung* von Evaluation „[...] entspricht der Vorstellung, dass aus Forschungserkenntnissen abgeleitete Empfehlungen mehr oder weniger direkt in Politikentscheidungen oder Praxishandeln umgesetzt werden“ (Müller, 2015, S. 29, vgl. auch Vedung, 2015, S. 190ff.; Weiss, 1998, S. 23f.). Dieser Nutzungstyp impliziert Funktionen wie die *Kontroll-, Steuerungs- und Managementfunktion* (Stockmann, 2007b; Caspari, 2004) sowie die *Optimierungs- und Entscheidungsfunktion* von Evaluationen (Döring & Bortz, 2016). Bei all diesen Funktionen wird Evaluation i.R. zu einem bestimmten Zweck instrumentalisiert und sie können in einem engen Zusammenhang zueinanderstehen: Evaluationen *kontrollieren* zu bestimmten Zeitpunkten, ob und inwieweit zuvor formulierte Zielsetzungen erreicht wurden oder Beteiligte ihren Aufgaben und Verpflichtungen nachkommen und ihre Qualifikationen oder Kompetenzen ausreichen. Abhängig von diesem *Soll-Ist-Vergleich* können *Entscheidungen* getroffen werden, bspw. über die Implementierung oder Weiterführung einer Maßnahme. Außerdem sind entsprechende *Steuerungsmaßnahmen* möglich, bspw. inwiefern ein Programm *verbessert* werden kann (*Optimierungsfunktion*) (vgl. Döring & Bortz, 2016, S. 987; Caspari, 2004, S. 24; Stockmann, 2017, S. 38). Im Zusammenhang mit M&E-Systemen wie dem *Logischen Modell* steht die von Vedung (2015, S. 201ff.) als *konstitutiv oder antizipatorisch* bezeichnete *Nutzung (constitutive, anticipatory use)*, bei der programmatische Inhalte und Aktivitäten an gewünschten Evaluationsergebnissen ausgerichtet werden. Auch die *rituelle Nutzung (ritual use)* (vgl. ebd., S. 200f.), womit routinierte und standardisierte Evaluationsprozesse im operationalen Geschäft einer Organisation gemeint sind, wie bspw. Schlussevaluationen bei der Beendigung einer Maßnahme, kann zu diesen im Managementbereich angesiedelten Funktionen und Nutzungsmöglichkeiten von Evaluation gezählt werden.

Von einigen Autor*innen weniger beachtet (vgl. Vedung, 2015; Weiss, 1998), von anderen mehr (Caspari, 2004; Döring & Bortz, 2016), ist die *Lern- und Dialogfunktion* von Evaluationen, die Stockmann (2007b, S. 38) auch unter *Entwicklung* fasst. Nur wenn Ergebnisse auf verschiedenen Ebenen transparent zwischen verschiedenen Stakeholdern

kommuniziert werden, kann Evaluation diesen Funktionen gerecht werden und eine Weiterentwicklung stattfinden. Dies beinhaltet auch die kritische Auseinandersetzung mit Ursachen für Erfolg und Misserfolg, wie erfolgreich eine Zusammenarbeit verläuft und welche Konsequenzen daraus für das weitere Vorgehen resultieren. Diese Evaluationsfunktion spielt bei formativ-gestaltenden Evaluationen eine besonders große Rolle (Döring & Bortz, 2016; Caspari, 2004).

Die *Legitimitätsfunktion* oder die *legitimierende Nutzung (legitimizing use)* von Evaluation ist in allen Funktions- und Nutzungszuschreibungen zu finden. Sowohl die Durchführung der Evaluation selbst als auch die Ergebnisse sollen eine Intervention nach außen hin legitimieren, bspw. wenn der Einsatz von öffentlichen Geldern gerechtfertigt werden soll (*Rechenschaftslegung, accountability*) (Döring & Bortz, 2016; Caspari, 2004; Stockmann, 2007b; 2017; Vedung, 2015; Weiss, 1998). In diesem Zusammenhang haben Evaluationen häufig auch eine *taktische* oder *strategische Funktion* oder werden dementsprechend genutzt (*tactical use*).¹¹²

Unabhängig, von wie vielen Funktionen ausgegangen wird, sind alle Evaluationsfunktionen eng miteinander verbunden: Ohne einen Erkenntnisgewinn kann keine andere Funktion erfüllt werden und andersherum werden mit jeder Funktion immer auch Erkenntnisse gewonnen. Die Steuerung von bspw. Projekten und Programmen steht nicht nur in einem Zusammenhang mit der Kontroll- oder Optimierungsfunktion, sondern auch mit der Entwicklung, die auch Lern- und Dialogfunktionen beinhaltet. Demnach stellt die Abbildung der Funktionen oder Nutzen von Evaluationen inklusive ihrer Beziehungen zueinander eine Herausforderung dar. Die folgende Abbildung sammelt die gängigsten Bezeichnungen und systematisiert an den von Müller (2015, S. 29) drei aufgeführten hauptsächlichen Verwendungstypen:

¹¹² Bspw. nutzen Politiker*innen Evaluationen als „dekorative Symbole für eine moderne Politik“, ohne dabei ernsthaft an Ergebnissen interessiert zu sein (Stockmann, 2006, S. 69) oder als „Entschuldigung für Untätigkeit“, wenn bspw. noch keine Entscheidung aufgrund fehlender Evaluationsergebnisse gefällt werden kann oder weil Ergebnisse als unklar oder unzureichend bezeichnet werden (Vedung, 2015, S. 199f.). Vedung (ebd.) sieht einen Unterschied zwischen der legitimierenden und taktischen Nutzung: „In tactical use, evaluation is an excuse for in action, whereas in legitimizing use, evaluation is employed to support extant substantive positions.“

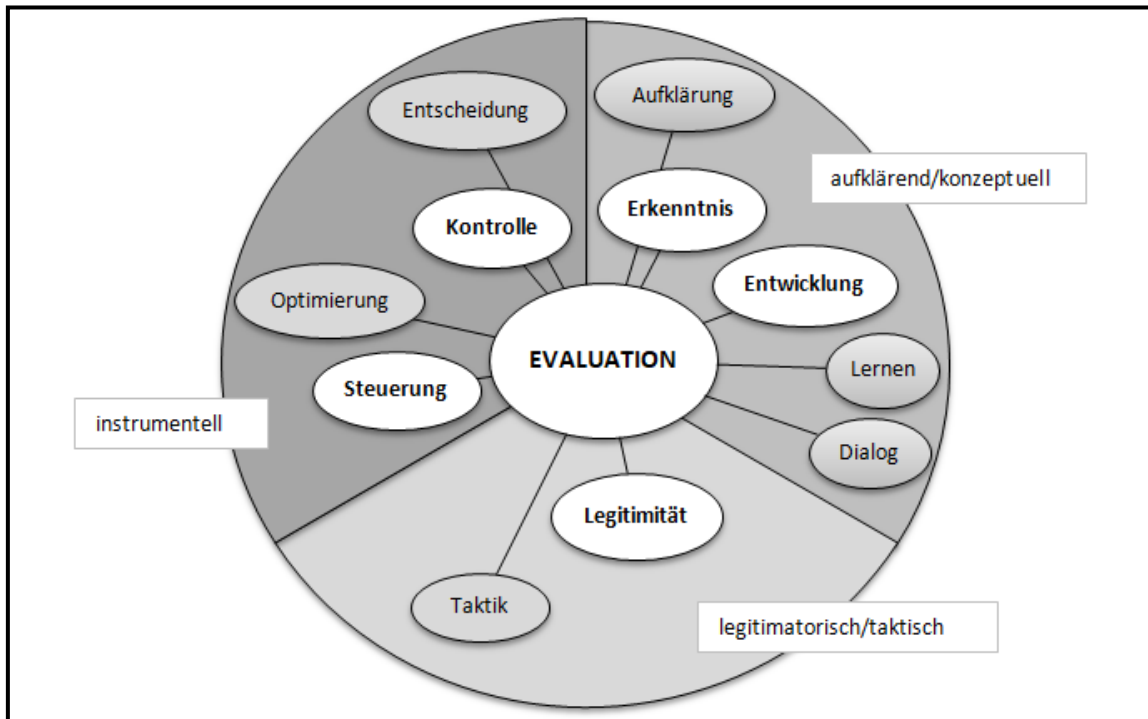


Abb. 7: Funktionen/Nutzen von Evaluationen (eigene Darstellung in Anlehnung an Caspari, 2004; Döring & Bortz, 2016; Müller, 2015; Stockmann, 2007b; 2017; Vedung, 2015, Weiss, 1998)

Eine große Nähe zur Bestimmung von Funktion und dem Nutzen von Evaluation hinsichtlich der Steuerungsfähigkeit und Legitimität von Entscheidungen hat der Begriff der *Evidenzbasierung* in der Politik und Praxis: Davies definiert *evidence-based policy and practice* „as an approach that helps people to make well-informed decisions about policies, programs and projects by putting the best available evidence from research at the heart of policy development and implementation“ (Davies, 2004, S. 3; zitiert nach Spiel & Schober, 2015, S. 41).

3.1.4 Zur Qualität von Evaluationen

Durch den stetigen Anstieg an Evaluationen ist die Qualitätsdebatte in der Evaluationsforschung sehr präsent und zugleich „schillernd“ (vgl. Caspari, 2015, S. 143; Widmer, 2000, S. 77). Auf der einen Seite geht es Evaluierenden darum, angemessene Untersuchungsziele zu stecken und entsprechende Forschungsmethoden zu entwickeln. Einschätzungen und Beurteilungen sowie Zusammenfassungen und Verallgemeinerungen von Studien werden dann möglich, sofern sie wissenschaftliche Standards erfüllen. Auf der anderen Seite wiederum sind Evaluationen immer an einen von außen an sie herangetragenen Zweck gebunden, die seitens der Auftraggebenden in sog. *Terms of Reference (TOR)* festgehalten werden und Eckdaten der Evaluation bestimmen (u.a. Untersuchungsgegenstand, Fragestellungen, möglicherweise Methoden, zeitlicher Rahmen etc.). Erfolgsfaktoren sind durch die Programmpraxis vorgegeben, die z.T. konträr zu den Erkenntniszielen und normativen Maßstäben aus der Wissenschaft stehen (Schemme, 2003; Stockmann & Meyer, 2017b; Wollmann, 2017). Die Evaluationsstandards *Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit* der Gesellschaft für Evaluation können i.R. dieses Spannungsverhältnisses als ‚Richtschnur‘ für „gute Evaluationen“ (DeGEval, 2016, S. 5) und als „Leitlinie in der Evaluationspraxis“ (bspw. in der Planung und Umsetzung) (Widmer, 2000, S. 94) dienen. Sie machen zudem die *Qualität von Evaluationen beurteilbar* (DeGEval, 2016; vgl. Widmer, 2000, S. ff.).¹¹³

Tab. 4: Standards für Evaluation (DeGEval, 2016, S. 18ff.)

Nützlichkeit
<p>N1 – Identifizierung der Beteiligten und Betroffenen Die am Evaluationsgegenstand oder an der Evaluation Beteiligten sowie die von Evaluationsgegenstand oder Evaluation Betroffenen sollen vorab identifiziert werden, damit deren Interessen und Informationsbedürfnisse geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können.</p>
<p>N2 – Klärung der Evaluationszwecke Es soll deutlich bestimmt sein, welche Zwecke mit der Evaluation verfolgt werden, so dass die Beteiligten und Betroffenen Position dazu beziehen und die Evaluierenden einen klaren Arbeitsauftrag verfolgen können.</p>
<p>N3 – Kompetenz und Glaubwürdigkeit des Evaluierenden Wer Evaluationen durchführt, soll fachlich und methodisch kompetent sein, damit für die Evaluation und ihre Ergebnisse ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.</p>
<p>N4 – Auswahl und Umfang der Informationen</p>

¹¹³ Weitere Orientierung bietet u.a. die Evaluationsstandards der Schweizer Evaluationsgesellschaft SEVAL (2016), die *Guiding Principles for Evaluators* der American Evaluation Association (AEA, 2018). Zur Entstehung der Standards s. Widmer (2000, S. 86f.).

<p>Auswahl und Umfang der erfassten Informationen sollen die adäquate Beantwortung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und die Informationsbedürfnisse der Auftraggebenden und weiterer Beteiligter und Betroffener berücksichtigen.</p>
<p>N5 – Transparenz von Werthaltungen Werthaltungen der Beteiligten und Betroffenen, die sich in deren Perspektiven und Annahmen manifestieren und einen Einfluss auf die Evaluation und Interpretation ihrer Ergebnisse haben, sollten transparent dokumentiert werden, um Evaluationsergebnisse besser einordnen zu können.</p>
<p>N6 – Vollständigkeit und Klarheit der Berichterstattung Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen und für ihre Adressatinnen und Adressaten verständlich und nachvollziehbar sein.</p>
<p>N7 – Rechtzeitigkeit der Evaluation Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungs-, Verbesserungs- oder sonstige Nutzungsprozesse einfließen können.</p>
<p>N8 – Nutzung und Nutzen der Evaluation Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation mitzutragen und ihre Ergebnisse zu nutzen.</p>
<p>Durchführbarkeit</p>
<p>D1 – Angemessene Verfahren Evaluationsverfahren, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, sollen so gewählt werden, dass einerseits die Evaluation professionell und den Erfordernissen entsprechend umgesetzt wird und andererseits der Aufwand für die Beteiligten und Betroffenen in einem adäquaten Verhältnis zum intendierten Nutzen der Evaluation gehalten wird.</p>
<p>D2 – Diplomatisches Vorgehen Evaluationen sollen so geplant und durchgeführt werden, dass eine möglichst hohe Akzeptanz der verschiedenen Beteiligten und Betroffenen in Bezug auf Vorgehen und Ergebnisse der Evaluation erreicht werden kann.</p>
<p>D3 – Effizienz von Evaluation Der Aufwand für Evaluation soll in einem angemessenen Verhältnis zum Nutzen der Evaluation stehen.</p>
<p>Fairness</p>
<p>F1 – Formale Vereinbarungen Die Rechte und Pflichten der an einer Evaluation beteiligten Parteien (was, wie, von wem, wann getan werden soll und darf) sollen schriftlich festgehalten werden.</p>
<p>F2 – Schutz individueller Rechte Evaluationen sollen so geplant und durchgeführt werden, dass Rechte, Sicherheit und Würde der in eine Evaluation einbezogenen Personen geschützt sind.</p>
<p>F3 – Umfassende und faire Prüfung Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst fair und umfassend prüfen und darstellen.</p>
<p>F4 – Unparteiische Durchführung und Berichterstattung Die Evaluation soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation beachten. Der gesamte Evaluationsprozess sowie die Evaluationsberichte sollen die unparteiische Position der Evaluierenden erkennen lassen.</p>
<p>F5 – Offenlegung von Ergebnissen und Berichten Evaluationsergebnisse und -berichte sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden.</p>
<p>Genauigkeit</p>
<p>G1 – Beschreibung des Evaluationsgegenstandes Sowohl das Konzept des Evaluationsgegenstandes als auch seine Umsetzung sollen genau und umfassend beschrieben und dokumentiert werden.</p>
<p>G2 – Kontextanalyse Der Kontext des Evaluationsgegenstandes soll ausreichend umfassend und detailliert analysiert sowie bei der Interpretation von Ergebnissen berücksichtigt werden.</p>
<p>G3 – Beschreibung von Zwecken und Vorgehen</p>

Zwecke, Fragestellungen und Vorgehen der Evaluation, einschließlich der angewandten Methoden, sollen so genau dokumentiert und beschrieben werden, dass sie nachvollzogen und beurteilt werden können.
G4 – Angabe von Informationsquellen Die i.R. einer Evaluation genutzten Informationsquellen sollen hinreichend genau dokumentiert werden, damit die Verlässlichkeit und Angemessenheit der Informationen eingeschätzt werden können.
G5 – Valide und reliable Informationen Erhebungsverfahren und Datenquellen sollen so gewählt werden, dass die Zuverlässigkeit der gewonnenen Daten und ihre Gültigkeit bezogen auf die Beantwortung der Evaluationsfragestellungen nach fachlichen Maßstäben sichergestellt sind. Die fachlichen Maßstäbe sollen sich an den Gütekriterien der empirischen Forschung orientieren.
G6 – Systematische Fehlerprüfung Die in einer Evaluation gesammelten, aufbereiteten, analysierten und präsentierten Informationen sollen systematisch auf Fehler geprüft werden.
G7 – Angemessene Analyse qualitativer und quantitativer Informationen Qualitative und quantitative Informationen einer Evaluation sollen nach fachlichen Maßstäben angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation beantwortet werden können.
G8 – Begründete Bewertungen und Schlussfolgerungen Die in einer Evaluation getroffenen wertenden Aussagen sollen auf expliziten Kriterien und Zielwerten basieren. Schlussfolgerungen sollen ausdrücklich und auf Grundlage der erhobenen und analysierten Daten begründet werden, damit sie nachvollzogen und beurteilt werden können.
G9 – Meta-Evaluation Meta-Evaluationen evaluieren Evaluationen. Um dies zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert, archiviert und soweit wie möglich zugänglich gemacht werden.

Widmer (2000) weist darauf hin, dass die Standards *Maximalansprüche* formulieren und es in der Praxis kaum möglich sein wird, alle 25 Einzelstandards zu erfüllen. „Trotzdem sollten alle Beteiligten [...] anstreben, die Standards so weit wie möglich zu berücksichtigen“ (ebd., S. 92).

Um die Qualität von Evaluationsprozessen und -ergebnissen zu überprüfen und sicherzustellen, plädieren Evaluationsforscher für mehr *Meta-Evaluationen* (Caspari, 2015), die Bestandteil der Standards sind (vgl. Tab. 4, G9). Die Standards selbst können bei Meta-Evaluationen als *Checkliste* eingesetzt werden (vgl. Caspari, 2015).

3.2 Evaluationsgegenstände und -schwerpunkte

Da unzählig vielfältige Gegenstände evaluiert werden können, ist die Bandbreite an *Evaluationsgegenständen*, auch als *Evaluationsobjekte* oder *Evaluanden* bezeichnet, sehr groß. Evaluiert werden können bspw. Vorhaben, Forschung, Gesetze, Einrichtungen, Institutionen, Organisationen, (Interventions-, Präventions-) Maßnahmen, Personen, Politikfelder, Produkte, Projekte, Programme, Strategien, Systeme und Lehre (vgl. Döring & Bortz, 2016, S. 979f.). Im Bereich der sozialwissenschaftlichen Evaluationsforschung wird häufig direkt von *Programmevaluation* (engl.: *programme evaluation*, vgl. bspw. Rossi et al., 2004; Mouton, 2014) gesprochen, da sie den größten Bereich i.R. der Evaluationsforschung ausmacht und besonders wichtig ist (vgl. Döring & Bortz, 2016, S. 980; Scriven, 2008, S. 285). „Evaluation stellt ein zentrales Instrument für die Planung, Implementation und Wirkungsmessung von Politiken, Strategien, Programmen, Projekten und Maßnahmen staatlicher und zivilgesellschaftlicher Akteure dar“ (Stockmann, 2018, S. 363). Sie wird genutzt, um die „[...] Effektivität, [...] Effizienz und [die] Wirksamkeit von sozialen und politischen Maßnahmen, Projekten [und] Programmen [...]“ zu steigern (Stockmann & Meyer, 2017a, S. 64). Da die Begriffe *Projekt*, *Programm*, *Vorhaben* oder *Maßnahme* in der für diese Arbeit relevanten Literatur oft unterschiedlich oder auch synonym verwendet werden, soll an dieser Stelle eine kurze Einordnung erfolgen.

3.2.1 Definitionen: Projekte, Programme, Vorhaben und Maßnahmen

Im Evaluationskontext unterscheidet Scriven (2008, S. 285f.) zwischen *Projekten* als „(...) time-bound efforts, often within a program“ und *Programmen* als „the general effort that marshals staff and projects towards some (often poorly) defined and funded goals“, wohingegen Stockmann (2006, S. 98) letztendlich zwischen Projekt und Programm nicht unterscheidet: „Instrumentell betrachtet handelt es sich bei Projekten und Programmen um *Maßnahmenbündel* zur Erreichung festgelegter Planziele, mit deren Hilfe Innovationen innerhalb sozialer Systeme eingeleitet werden sollen.“ Döring und Bortz (2016) nutzen den Begriff *Maßnahmen* stellvertretend für das gesamte Spektrum möglicher Evaluationsgegenstände. Als *Entwicklungsmaßnahmen* bezeichnet das BMZ (2008) eine übergeordnete Durchführungsform der dt. EZ für die bilaterale finanzielle und techni-

sche Zusammenarbeit mit Kooperationspartnern. Oberndörfer et al. (2010, S. 240) sprechen von „jegliche[r] abgrenzbare[n] Maßnahme (Projekt, Programm, Politik usw.) zur Erzielung einer entwicklungspolitischen Wirkung“. Projekte und Programme sind wichtige Formen dieser Entwicklungsmaßnahmen. *Projekte* werden als „[...] thematisch bzw. sektoral, räumlich, wirtschaftlich und zeitlich abgrenzbare Maßnahmen mit zielorientierter Aufgabenstellung“ definiert. Eine Bündelung von Projekten wird für sinnvoll erachtet, wenn sie inhaltliche, organisatorische oder räumliche Gemeinsamkeiten aufweisen: „*Programme* sind Entwicklungsmaßnahmen, die in der Regel aus mehreren Komponenten bestehen, auf verschiedenen Interventionsebenen agieren und denen ein einheitliches Konzept zugrunde liegt“ (BMZ, 2008, S.12). Die GIZ (2015a, S. 3) verwendet den Begriff *Vorhaben* für Programme und Projekte, „die sich zum Ziel gesetzt haben, innerhalb eines definierten Zeitrahmens gesellschaftliche, politische oder wirtschaftliche Veränderungen zu unterstützen“. Bei sog. *Sektorvorhaben* handelt es sich um überregionale Maßnahmen der technischen Zusammenarbeit, die „[...] durchführungsnahe Methoden mit dem Ziel [erarbeiten], die entwicklungspolitischen Wirkungen in einem Sektor¹¹⁴ bzw. zu einem sektoralen Thema zu erhöhen. Sie umfassen i.d.R. *Pilotmaßnahmen* in Kooperationsländern.“ (GIZ, 2018, 2. Juli).

Projekte, Programme und andere politische Interventionsmaßnahmen sind in der Regel an *Politiken* (*policies*) ausgerichtet. Die inhaltliche Dimension des Politikbegriffs beschreibt eine in sich geschlossene Handlungsstrategie in Bezug auf ein spezifisches Themen- oder Problemfeld.¹¹⁵ Programme und Projekte bilden hierbei Kernelemente solcher Strategien, um übergeordnete politische Zielsetzungen zu erreichen (vgl. Bussmann

¹¹⁴ Bspw. Gesundheit, Bildung oder Landwirtschaft. Die GIZ listet i.R. ihrer Projektdarstellung fast 40 unterschiedliche Sektoren (vgl., GIZ, 2018). Auch *Jugendschutz* oder *Fluchtursachenbekämpfung* werden als Sektoren begriffen, die verschiedene Sektoren betreffen. Die Begrifflichkeiten *Sektoren*, *Politikfelder*, *Teilbereiche* oder *-systeme der Gesellschaft*, *gesellschaftliche Handlungsfelder* und *Kooperationssysteme auf Dauer* werden von der GIZ synonym genutzt und bezeichnen „[...] stets das *Umfeld*, auf das sich die Ziele eines Vorhabens beziehen und in dem eine Veränderung angestrebt wird“ (GIZ, 2015a, S. 3f.).

¹¹⁵ I.d.R. werden i.R. des Politik-Begriffs drei Dimensionen unterschieden: (1) *Policy* beschreibt die *inhaltliche* Dimension, d.h. wie gesellschaftliche Probleme durch das politische System wahrgenommen, verarbeitet und welche ziel- oder zweckgerichteten Maßnahmen angewendet werden; (2) *Politics* bezieht sich auf die *prozessualen* Aspekte, bspw. den Verlauf politischer Prozesse; (3) *Polity* beschreibt die *formalen/strukturellen* Aspekte von Politik, d.h. die Form, in der sie abläuft (Böhret, Jann & Kronenwett-Löhrlein, 1988, S. 3ff.).

et al., 1997, S. 83), wie bspw. die Nachhaltigen Entwicklungsziele (SDGs) der Vereinten Nationen (vgl. BMZ, 2018).

In dieser Arbeit umfasst der Begriff der *Maßnahmen* Projekte und Programme der EZ. Zwischen Programmen und Projekten wird unterschieden, wenn dies als notwendig erachtet wird. Des Weiteren wird im Kontext der SfE-Maßnahmen der GIZ von *Regional- und Sektorvorhaben* (RV und SV) die Rede sein, die verschiedene Maßnahmen in unterschiedlichen Ländern umsetzen.

3.2.2 Zum Verlauf von Maßnahmen

Maßnahmen durchlaufen unterschiedliche Phasen, die mit denen eines *Lebensverlaufs* gleichgesetzt werden können. Stockmann (2006, S. 109f) überträgt konzeptuelle Annahmen der Lebenslaufforschung¹¹⁶ (vgl. Huinink, 1995, S. 154f.) auf die Programmentwicklung¹¹⁷. Wie ein Lebensverlauf ist ein Programmverlauf ein *selbstreferenzieller Prozess*, d.h. es gibt unterschiedliche Phasen, deren Abfolge in einem kausalen Zusammenhang stehen. Er ist ein *multidimensionaler Prozess*, d.h. er setzt sich aus mehreren Bereichen zusammen (bspw. strategische Entwicklung, Organisationsentwicklung, konzeptionelle Entwicklung etc.), die sich wechselseitig aufeinander beziehen und gegenseitig beeinflussen. Zudem ist er *Teil eines gesellschaftlichen Mehrebenenprozesses*, d.h. andere Elemente (bspw. andere Maßnahmen; beteiligte Akteur*innen; vorhandene gesellschaftliche, institutionelle, organisatorische Rahmenbedingungen; soziale, regionale Kontexte und Veränderungen) haben strukturierende Einflüsse auf Programmverläufe. Anders als Lebensverläufe werden Maßnahmen von Beginn an häufig bis in jedes Detail geplant, danach umgesetzt und finden einen terminierten Abschluss. Hierbei können drei Hauptphasen unterschieden werden (Stockmann, 2006, S. 111): Die *Planungsphase*, *Implementationsphase* (Umsetzung) und die *ex-post Wirkungsphase* nach Beendigung eines Programms. t1 in Abbildung 8 markiert den Beginn eines Programmverlaufs, bspw. durch die Formulierung einer *Programmidee*. Daraufhin durchläuft ein Programm seine

¹¹⁶ „Die Lebenslaufsforschung ist eine spezifische Form der Sozialstrukturanalyse“ (Huinink & Mayer, 1995, S. 8). Sie beschäftigt sich mit der „[...] Untersuchung sozialer Prozesse, die sich über den gesamten Lebensverlauf oder wesentliche Teile davon erstrecken [...]“ und die im Kontext des institutionellen Wandels und besonderer historischer Ereignisse betrachtet werden (Mayer, 1990, S. 9).

¹¹⁷ Stockmann (2006) nutzt den Begriff *Programme*, der in den folgenden Abschnitten übernommen wird.

verschiedenen Phasen, wie die *Programmprüfung* (t2), die *Angebotsformulierung* (t3), die *Auftragserteilung* (t4), verschiedene Phasen der *Programmdurchführung* (t5a-t5n) bis hin zum *Förderende* (tF) und die *Zeit nach der Förderung* (tNF). Die Phasen lassen sich voneinander abgrenzen und können mittels im Prozess produzierter Daten, bspw. in Angeboten, Operationsplänen, Fortschrittsberichten, Evaluations- und Abschlussberichten etc. analysiert werden (ebd.).

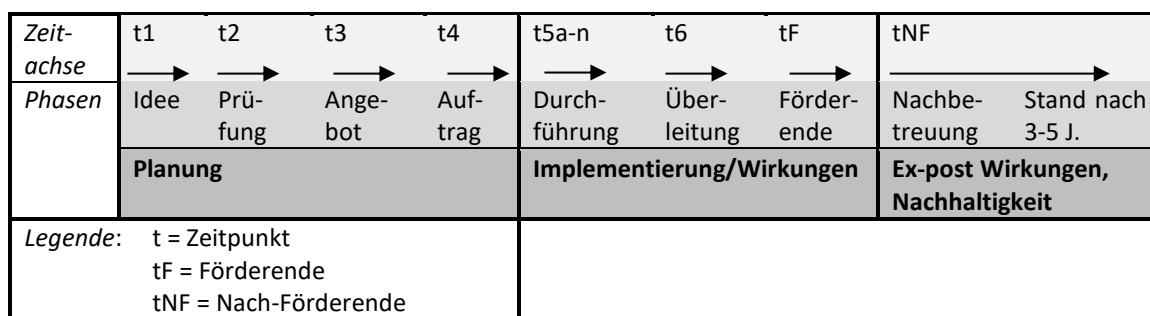


Abb. 8: Lebensverlaufsmodell einer Maßnahme (Stockmann, 2006, S. 111)

Stockmann (2006, S. 111) sieht den heuristischen Vorteil der Lebensverlaufsperspektive vor allem bzgl. zweier Aspekte: Erstens wird „[...] die *Nachförderphase*, in der sich die Nachhaltigkeit eines Programms zeigt“, als integraler Bestandteil betrachtet und zweitens wird „[...] die *kausale Verkettung* der einzelnen Phasen hervor[gehoben]“, die verdeutlicht, dass die *Nachhaltigkeit* bereits schon durch die Planung beeinflusst wird „[...] und dass die während der Förderlaufzeit geschaffenen materiellen und immateriellen Strukturen das Fundament für die *langfristigen Programmwirkungen* bilden.“

3.2.3 Evaluation von Wirkungen und Nachhaltigkeit von Maßnahmen

Auftraggebende, Durchführungsorganisationen und im weiteren Sinne auch Steuerzahlende möchten „wissen, was wirkt“¹¹⁸. D.h. sie möchten wissen, was Maßnahmen nachhaltig bewirken und warum dies der Fall ist oder nicht. Demnach ist das Interesse an Evaluationen, die konkrete Aussagen diesbezüglich treffen können, groß, vor allem wenn es um Politikgestaltung und um Wissensgenerierung geht (vgl. Mark & Henry, 2007). Doch die Darstellung solcher *Ursache-Wirkungszusammenhänge* stellt häufig eine Herausforderung dar. In enger Verbindung mit langfristigen *Wirkungen* steht die

¹¹⁸ Zahlreiche GIZ-Broschüren wurden unter diesem Motto veröffentlicht (vgl. u.a. Karkoschka, 2017; Karkoschka, Milcher, Müller –Gerold & Bendtfeld-Huthmann, 2017; GIZ, 2017; GIZ Stabsstelle Evaluierung, 2018a-e)

Nachhaltigkeit von Maßnahmen, d.h. was ein Projekt oder Programm auch nach Beendigung seiner Laufzeit nachhaltig bewirkt. Beide Komponenten sollen im Folgenden näher betrachtet und Evaluationsmöglichkeiten eruiert werden.

Wirkungen: Modelle, Designs, Methoden und Störfaktoren

Analytisch lassen sich Wirkungen anhand dreier Dimensionen bestimmen (Stockmann, 2006, S. 103). Wirkungen können (1) *Strukturen* (z.B. von Organisationen oder gesellschaftlichen Teilsystemen), *Prozesse* und/oder individuelle *Verhaltensweisen* betreffen; (2) geplant (*intendiert*) oder ungeplant (*nicht-intendiert*) auftreten; (3) *positiv* oder *negativ* sein, d.h. Maßnahmenziele unterstützend oder ihnen zuwiderlaufend. Zudem unterscheiden sie sich in ihrer *Art* (ökonomisch, sozial, ökologisch, kulturell, politisch), in ihrer *zeitlichen Dauer* (kurz-, mittel-, langfristig), in ihrer *Ebene des Auftretens* (*makro/Gesellschaft, meso/Organisation oder Gruppe, mikro/Individuen*) und hinsichtlich ihres Wegs (*direkt, indirekt*). Des Weiteren kann zwischen *internen*, bei der durchführenden Organisation ausgelösten Wirkungen, und *externen* Wirkungen unterschieden werden, die sich auf das Umfeld der Maßnahme beziehen (bspw. Zielgruppen, andere Politikfelder) (Reade, 2008). Die Wirkungsmöglichkeiten eines Projekts oder Programms können einerseits durch *interne Faktoren* einer (Durchführungs-)Organisation (Ziele, Mitarbeitende, Organisationsstruktur, Ressourcen, Technologien) und andererseits durch *externe (Umwelt-)Faktoren*, d.h. andere soziale, politische, kulturelle, ökonomische Systeme, die das Programm umgeben, beeinflusst werden (Stockmann, 2006).

Abbildung 9 veranschaulicht anhand des *Logischen Modells* die *Zielebenen von Wirkungsevaluationen*.

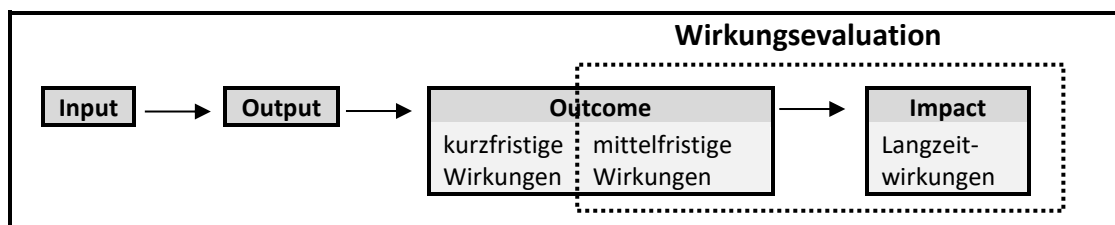


Abb. 9: Logisches Modell mit Zielebenen von Wirkungsevaluationen (eigene Darstellung nach Caspari & Barbu, 2008, S. 5)

Häufig werden i.R. von Maßnahmen lediglich die Zielerreichung oder die kurzfristigen Wirkungen auf die Zielgruppe untersucht (*Outcomes*). Dabei interessieren bei Wirkungsevaluationen vielmehr die mittelfristigen, direkten sowie die langfristigen Wirkungen (*Impacts*) einer Maßnahme, die *nur über ex-post Evaluationen* zu ermitteln sind (Caspari & Barbu, 2008). Die Erarbeitung eines solchen logischen Modells für Maßnahmen, das eine differenzierte *Ursache-Wirkungszuschreibung* beinhaltet und mit *theoriebasierten Wirkungshypothesen* gestützt ist, stellt die erste wichtige Voraussetzung für Wirkungsevaluationen dar, da somit ein sog. *theoriebasierter Ansatz* zur Wirkungsmessung angewendet wird (vgl. Stockmann, 2006, S. 105, S. 67). Evaluationen, die Ergebnisse über die Wirkungen geben, aber keine Angaben machen, *warum* eine Maßnahme die erwarteten Wirkungen gezeigt hat oder nicht, hinterlassen eine *black box* mit fehlenden Informationen in der Kausalkette einer Wirkung (Caspari & Barbu, 2008; Reade, 2008).

Wirkungsevaluationen stehen vor zahlreichen *methodischen und praktischen Herausforderungen*. Zum einen sollen sie Aufschluss über möglichst *alle Wirkungen* geben, sowohl den seitens des Projekts oder Programms intendierten als auch den nicht-intendierten Veränderungen. Zum anderen sollen sie die Wirkungen *kausal zuschreiben* können, d.h. es geht darum, die *Nettowirkungen* einer Maßnahme zu isolieren.¹¹⁹ Es handelt sich hierbei um die *Berücksichtigung des Kontrafaktischen*, d.h. Wirkungsevaluationen analysieren Wirkungen einer Maßnahme, die nicht beobachtbar gewesen wären, wenn die Maßnahme nicht durchgeführt worden wäre. In diesem Zusammenhang wird vom *Zuordnungsproblem* oder der *Zuordnungslücke (attribution gap/problem)* gesprochen (Caspari & Barbu, 2008; Oberndörfer et al, 2010; Reade, 2008). Wenn eine Evaluation „[...] Wirkungen möglichst eindeutig Interventionen zuordnen kann (*Attribution*), wenn der Nachweis gelingt, dass ohne Intervention die Wirkungen nicht aufgetreten wären (*Berücksichtigung des Kontrafaktischen*), und wenn die interne Validität (*Designqualität*) sowie die externe Validität (*Generalisierbarkeit der Befunde*) gegeben sind“, wird sie als *rigoros* bezeichnet (GIZ Stabsstelle Evaluierung, 2018a, S. 16). Neben dem Zuordnungsproblem existieren noch weitere methodische und praktische Herausforderungen, die

¹¹⁹ *Kausalmechanismen* generieren die beobachteten Wirkungen und beschreiben, was zwischen Ursache und Wirkung passiert. Durch die Analyse von Kausalmechanismen kann die kausale schwarze Box geöffnet und die Frage nach dem *Wie* und *Warum* beantwortet werden (Schmitt, 2018).

von Caspari und Barbu (2008) als *Störfaktoren* bezeichnet werden und die „[...] einen entscheidenden Einfluss auf die *Qualität* der Ergebnisse [haben]“ (ebd., S. 16). Dazu gehören bspw. *zeitliche Umstände* (Zeitpunkt(e) der Datenerhebung, Verzögerung etc.), Auswahlverzerrungen (*selection bias*) und sog. *Übertragungseffekte*.¹²⁰ Die teilweise komplexen Voraussetzungen, die bei einer Wirkungsevaluation im *Idealfall* gegeben sein sollten, können *in der Praxis* nicht immer erfüllt werden, so wie bspw. die Erhebung von Baseline-Daten oder der Einbezug von Kontrollgruppen (ebd.). Oberndörfer et al. (2010, S. 31) sehen auch in der *Wahl eines geeigneten Verfahrens der Wirkungsanalyse*¹²¹ eine praktische Herausforderung und weisen in diesem Zusammenhang auf Interessenkonflikte bei den beteiligten Akteur*innen sowie benötigte zeitliche, personelle und finanzielle Ressourcen (Mittelaufwand) hin, die von Verfahren zu Verfahren sehr unterschiedlich sind.

Um Zuordnungsprobleme überwinden zu können, sind *spezielle und komplexe Datenerhebungsdesigns* notwendig. Das Forschungsdesign ist letztendlich entscheidend „für den Grad der Gewissheit, mit dem die Frage nach dem Zusammenhang zwischen [...] Ursache und Wirkung [...] beantwortet werden kann“ (Schnell, Hill & Esser, 2011, S.201). Stockmann (2006, S. 226ff.) beschreibt und veranschaulicht tabellarisch acht Designs, die von oben nach unten im Grad an Komplexität und Genauigkeit abnehmen.¹²² Sog. *Randomized Controlled Trials (RCT)* stehen an erster Stelle, das *Gutachtenmodell*, wobei Expert*innen die Netto-Wirkungen einer Maßnahme einschätzen sollen, als kostengünstiges, aber unzuverlässigstes und ungenauestes Design an letzter Stelle (ebd., S.

¹²⁰ Mit *Übertragungseffekt* ist gemeint, dass eine Maßnahme über die intendierte Zielregion hinauswirken kann, bspw. in Form eines Bauprojekts. Caspari und Barbu (2008, S. 15f.) erläutern alle Störeffekte im Detail. Schnell et al. (2011, S. 207ff.) tragen weitere Störeffekte zusammen, die bei der Nutzung von Methoden der empirischen Sozialforschung auftreten können. S. Definitionen von Verzerrungen bzw. *Bias* auch bei Scriven (2008, S. 67ff.).

¹²¹ Nach Oberndörfer et al. (2010, S. 242) „[werden] in einer *Wirkungsanalyse* [...] entwicklungspolitische Wirkungen sozialwissenschaftlich nachgewiesen bzw. plausibilisiert. In einer *Wirkungsevaluierung* werden die Wirkungen zusätzlich bewertet“. Im Text werden die Begriffe vornehmlich synonym genutzt, auch weil die Unterscheidung nach Oberndörfer et al. (2010) in weiterer Literatur keine Beachtung findet.

¹²² Stockmann (2006, 224ff.) erläutert alle Designs ausführlich und gibt einen Einblick in Vor- und Nachteile.

229).¹²³ Unabhängig der Vorteile von qualitativen (*Konstruktivismus*) und quantitativen Methoden (*Positivismus*), wird im Zuge von Wirkungsevaluationen vermehrt eine Kombination beider Methoden im Sinne einer *Triangulation/Methodenmix (Mixed Methods)* empfohlen, u.a. um keine *black box* in der Kausalkette der Wirkungen zu hinterlassen (Caspari & Barbu, 2008; Reade, 2008).

Nachhaltigkeitskonzepte und Nachhaltigkeitsevaluation

Die eine Definition von Nachhaltigkeit ist nicht existent. Mit der Begriffsunklarheit droht die Gefahr, „[...] dass das Konzept der nachhaltigen Entwicklung zu einer ‚Leerformel‘, einer ‚Worthülse‘, einem ‚politischen Schlagwort‘ verkommt“ (Stockmann, 2006, S. 136). Der ursprünglich aus der Forstwirtschaft des 19. Jahrhunderts stammende Begriff zur Beschreibung einer langfristigen, also nachhaltigen Waldnutzung (Goetze, 2002), wurde 1987 über die Vereinten Nationen im sog. *Brundtland-Report* eingeführt. Dieser Report etablierte erstmalig den Begriff von *sustainable development* und führte somit zwei wegweisende Aspekte von künftiger Entwicklungspolitik zusammen, die zuvor kaum Beachtung gefunden haben oder getrennt voneinander behandelt wurden (vgl. Bethge, Steurer & Tscherner, 2011, S. 26f.). Zudem wurde zum ersten Mal „[...] die politische Verknüpfung von *Ökologie* mit *Wirtschaft* und *Sozialem* generiert“ (Caspari, 2004, S. 46).¹²⁴ Demnach ist der Begriff der nachhaltigen Entwicklung „im Kontext der Diskussionen der internationalen Staatengemeinschaften entstanden“ (Stockmann, 2006, S. 135) und hat sich als Leitbild internationaler Agenden ausdifferenziert und weiterentwickelt.¹²⁵ In diesem Kontext ist der Begriff der *Nachhaltigkeit* (eng.: *sustainability*) durch die Definition der OECD (2009, S. 46) geprägt und meint das „Fortbestehen der positiven Effekte einer Entwicklungsmaßnahme nach Abschluss der EZ-Leistungen; Wahrschein-

¹²³ Die *experimentelle Untersuchung* wird in der quantitativen Sozialforschung häufig als der „Königsweg“ oder „Goldstandard“ kausaler Wirkungsevaluation bezeichnet (vgl. Caspari & Barbu, 2008; Mark & Henry, 2007; Müller, 2017; Neubert, 2010; Oberndörfer et al., 2010; Reade, 2008; Roxin, 2016; Scriven, 2008). Den Verlauf eines experimentellen Designs beschreibt Neubert (2010, S. 146) und Müller (2017, S. 207ff). Müller (2017) nennt zudem Herausforderungen und weitere Designs, wenn die Praxis keine RCTs zulässt.

¹²⁴ Eine Übersicht zu den Nachhaltigkeitsdimensionen gibt König (2011, S. 45ff.).

¹²⁵ Auf eine Darstellung der Vorgeschichte zum Nachhaltigkeitsdiskurs und auf eine ausführliche historische Abhandlung der Meilensteine wird an dieser Stelle verzichtet. Einen detaillierten Überblick bieten u.a. Bethge et al. (2011), Caspari (2004), Martinuzzi und Meyer (2017), aus OECD-Perspektive Strange und Bayley (2008).

lichkeit positiver Langzeiteffekte; Langfristige Widerstandsfähigkeit der Nettonutzeffekte gegenüber Risiken.“ Wie in anderen Definitionen im EZ-Kontext wird auch hier der Begriff der Nachhaltigkeit synonym mit *langfristigen Wirkungen* bzw. der *quantifizierten Dauerhaftigkeit* von Maßnahmen gleichgesetzt, was Evaluationsforscher*innen kritisieren. Dem Begriff immanente Aussagen wie zur *Qualität von Wirkungen* wird keine Aufmerksamkeit geschenkt (vgl. Caspari, 2004, S. 54f.; Stockmann, 2006, S. 138) und die Wichtigkeit von Nachhaltigkeitsevaluation im Sinne transformativer Evaluation vernachlässigt (Patton, 2019; Schröter 2020). Entwickelte Evaluationsansätze versuchen diesem Defizit entgegenzuwirken, haben sich bislang allerdings noch nicht in der Praxis etabliert.¹²⁶

So wie die Wirkungsmöglichkeiten einer Maßnahme, ist auch deren Nachhaltigkeit von unterschiedlichen Faktoren beeinflusst: Neben *internen Organisationsfaktoren* (u.a. die Stärke einer Organisation, der Integrationsgrad der Programme in der Organisation, Kompetenzen der Mitarbeitenden und Führungskräfte) sowie *Umweltfaktoren* (politisches, soziales, wirtschaftliches Umfeld; Grad der Community-Beteiligung) identifizieren Shediac-Rizkallah und Bone (1998) zudem *Design- und Implementierungsfaktoren*, womit Verhandlungsprozesse, verfügbare Ressourcen sowie CD-Maßnahmen zur Nachhaltigkeit gemeint sind, die letztendlich aber auch zu den internen Faktoren gezählt werden können.

¹²⁶ Schröters (2020) Konzept der *Evaluation von Nachhaltigkeit für Nachhaltigkeit* (*Evaluation of sustainability for sustainability*) beinhaltet eine umfangreiche Checkliste für Praktizierende und deckt zwei Perspektiven ab: zum einen die *Evaluation von Nachhaltigkeit*, d.h. wie dauerhaft besteht ein Evaluationsgegenstand (bspw. Projekte und Programme) und zum anderen die *Evaluation für Nachhaltigkeit*, d.h. welchen Beitrag der Evaluationsgegenstand zur nachhaltigen Entwicklung leistet. Die *Nachhaltigkeitskonzeption* nach Stockmann (2006) beinhaltet vier Dimensionen, die über die Dauerhaftigkeit von Projekten hinausgehen, und einen *additiven Nachhaltigkeitsindex*, wodurch zusammenfassende Aussagen zur Nachhaltigkeit einer Maßnahme möglich sind (vgl. Caspari, 2004). Letzterer Ansatz wird am Ende der Arbeit detaillierter erläutert (s. Kap. 7.3.1).

3.3 Akteur*innen im Evaluationsprozess

Die Entwicklung von Evaluation hängt eng mit der Nachfrage seitens staatlicher, aber auch nichtstaatlicher, Stellen und Organisationen, also der Auftraggebenden, ab (Stockmann, 2017; 2018). Neben den Auftraggebenden und Auftragnehmenden, d.h. denjenigen, die die Evaluation durchführen (Evaluierende, Evaluationsteam), sind noch weitere Akteur*innen am Evaluationsprozess beteiligt. Die Beteiligten stehen in unterschiedlichen *sozialen Beziehungen* zueinander, die nicht außer Acht gelassen werden sollten: Sie haben bspw. Einfluss auf die Qualität und Effektivität einer Maßnahme (vgl. Abma, 2007, S. 185ff.). Im Folgenden werden zunächst unterschiedliche Evaluationsperspektiven veranschaulicht (interne, extern), im Anschluss alle beteiligten Stakeholdergruppen ermittelt und zuletzt die Rolle der Evaluierenden selbst ins Auge gefasst.

3.3.1 Rollenverteilung und institutionelle Zugehörigkeit

Mit Blick auf die Akteur*innen kann bzgl. der Rollenverteilung zwischen Evaluierenden und Evaluierten (*Fremd- und Selbstevaluation*) und der institutionellen Zugehörigkeit (*interne und externe Evaluation*) unterschieden werden, wobei die jeweiligen Perspektiven immer *Vor- und Nachteile* bergen (vgl. Brandt, 2007, S. 169ff.; Caspari, 2004, S. 32ff.; Döring & Bortz, 2016, S. 989f.; Stockmann, 2007b, S. 61f.; 2006, S. 263).

Evaluationen sind *intern*, wenn sie von derselben Organisation vorgenommen werden, die auch die Maßnahme durchführt. Wenn dabei die Abteilung sowohl die operative Implementierung als auch die Evaluation übernimmt, wird von einer *Selbstevaluation* gesprochen. Vorteile von sog. *in-house*-Evaluationen sind eine rasche und mit geringem Aufwand verbundene Durchführung, eine hohe Sachkenntnis und die Möglichkeit der unmittelbaren Umsetzung von Ergebnissen. Zudem kennen interne Evaluierende bereits beteiligte Akteur*innen und können auf vorhandene Kommunikationsstrukturen aufbauen. Nachteile können in einer geringen Methodenkompetenz, in der fehlenden Unabhängigkeit bzw. Distanz zum Untersuchungsgegenstand sowie in der daraus resultierenden ‚Betriebsblindheit‘ liegen. Vor allem bei Selbstevaluationen kann es zu *Rollenkonflikten* kommen, „da eine objektive Bewertung mit negativen Ergebnissen bzgl. des Programmerfolgs gleichbedeutend mit einer negativen Kritik an der eigenen Arbeit ist“ (Brandt, 2007, S. 170).

Externe (Fremd-)Evaluationen werden von Personen durchgeführt, die nicht dem Fördermittelgebenden oder der Durchführungsorganisation angehören. Für sie spricht eine hohe Unabhängigkeit, eine profunde Methodenkompetenz sowie eine hohe Glaubwürdigkeit bzgl. der Ergebnisse. Externe Evaluierende sehen sich allerdings auch mit der Herausforderung konfrontiert, eine Maßnahme weniger gut zu kennen und können auf Abwehrreaktionen bei den Evaluierten stoßen. Zudem werden möglicherweise Ergebnisse oder Empfehlungen weniger akzeptiert und im Anschluss umgesetzt. Zudem besteht die Gefahr, dass auf dem freien Markt durch Folgeaufträge zwischen der auftraggebenden Organisation und der evaluierenden Person *Abhängigkeiten* geschaffen werden können und die *Unabhängigkeit* in der Bewertung in Frage gestellt werden kann. Bspw. kann es zur Vernachlässigung von Perspektiven kommen (Fokus auf Auftraggebenden) oder kritische Bewertungen werden zurückgenommen (Brandt, 2007).

Die hier beschriebenen verschiedenen Möglichkeiten der institutionellen Einbettung von Evaluation sind idealtypisch. Im Einzelfall können sie zutreffen, müssen es aber nicht zwangsläufig. Wichtig ist, dass sie die Möglichkeiten und Schwierigkeiten von Evaluierenden in ihrem Arbeitsfeld veranschaulichen. Wenn interne und externe Evaluationen kombiniert werden, können durch beide Perspektiven *Synergien* geschaffen und Vorteile beider Vorgehensweisen genutzt werden (vgl. Stockmann, 2007b, S. 62).

3.3.2 *Beteiligte Stakeholdergruppen*

Da Maßnahmen meist eine Mehrebenenstruktur immanent ist (mehrere Projekte, Planung- und Umsetzungsebenen), bringen sie ein „komplexes Zusammenwirken vieler beteiligter Programmakteure mit sich“ (*Multiakteursstruktur*) (Heinze et al., 2019, S. 266). Demnach sind an einer Evaluation zahlreiche Stakeholdergruppen beteiligt oder davon betroffen (vgl. Brandt, 2007, S. 166): *Auftraggebende*, die die Evaluation in Auftrag geben; *Geldgebende*, die die Evaluation finanzieren (können identisch sein mit dem Auftraggebenden); *Geldgebende*, die die zu evaluierende(n) Maßnahme(n) finanzieren; *Programmmanager*innen und -mitarbeitende*, die für die Maßnahme auf strategischer und operativer Ebene verantwortlich sind; *Zielgruppen*, auf die das Vorhaben ausgerichtet ist; das *Evaluationsteam/Evaluierende* selbst und *sonstige Betroffene*, d.h. Personen o-

der Gruppen mit wenig oder keinem Einfluss auf die Evaluationsdurchführung, die jedoch mit dieser in Verbindung stehen (bspw. Personen aus dem weiteren Umfeld der Maßnahme; Programmkonkurrent*innen). Darüber hinaus sind häufig auch *lokale Partnerorganisationen* (bspw. NGOs) an Evaluationsprozessen beteiligt. Sie fungieren als sog. *Gatekeeper*, indem sie in erster Linie den *Zugang zu Zielgruppen* sichern und können darüber hinaus weitere wichtige Aufgaben/Funktionen übernehmen, bspw. die Übersetzung von Instrumenten oder die Datensammlung selbst (ebd.).

Die Aufzählung veranschaulicht, dass einzelne Gruppen mit sehr *unterschiedlichen Interessen, Perspektiven und Fragestellungen* an einer Evaluation direkt oder indirekt beteiligt sind. Bspw. kann es einem Geldgebenden vornehmlich darum gehen, ob eine Maßnahme kosteneffizient umgesetzt wird, während lokale Mitarbeitende vorrangig an der Weiterführung interessiert sind. Evaluierende müssen von daher in der Lage sein, mit diesen unterschiedlichen Interessensgruppen zu kommunizieren und ggf. zwischen ihnen zu vermitteln. Sie können die Planung und Durchführung von Evaluationen mehr oder weniger fördern oder gar behindern: „Aus der Grundproblematik verschiedener Interessenskonstellationen ergeben sich unweigerlich Spannungen, Belastungen und evtl. Konflikte, denen nur zum Teil durch die Evaluatoren entgegengewirkt werden kann“ (Brandt, 2007, S. 168; vgl. Döring & Bortz, 2016, S. 976; Fiege, 2012, S. 239f.). Die Akteur*innen befinden sich demnach in einen fortwährenden „hochgradig situativen und komplexen Aushandlungsprozess“ (Heinze et al., 2019, S. 268). Die in Kapitel 3.1.4 aufgeführten Evaluationsstandards können Interaktionen zwischen Evaluierenden und Stakeholdern „[...] so regeln, dass ‚respektvoll und fair‘ [s. *Fairness-Standards* DEGeval, 2016] miteinander umgegangen wird“ (Stockmann, 2006, S. 265). Des Weiteren sollten im Sinne der *Nützlichkeitsstandards* (vgl. DeGeval, 2006, S. 18f.) alle beteiligten Akteur*innen bereits zu Beginn ihre jeweiligen Interessen, Zielsetzungen und Bedenken bzgl. geplanter Evaluationsprozesse und -ergebnisse transparent formulieren. Dies kann potenziellen Interessenskonflikten sowie negativen Einstellungen gegenüber Evaluationsaktivitäten entgegenwirken (Brandt, 2007). Im weiteren Verlauf sollten sie sowohl beim Planungsprozess als auch bei den darauffolgenden Durchführungsschritten im besten Falle *aktiv* eingebunden, zumindest aber ausreichend informiert sein. Dadurch wird zum einen eine hohe *Transparenz*, zum anderen eine breite *Akzeptanz* geschaffen, was

wiederum zu einer erfolgreicherer Umsetzung von Forschungsprozessen, unverfälschteren Forschungsergebnissen und am Ende zur effektiveren Nutzung der Evaluationsergebnisse führt (vgl. ebd., S. 167; Döring & Bortz, 2016, S. 982; Schemme, 2003, S. 32). Auch Rossi et al. (2004, S. 50) empfehlen die Einbindung von Stakeholdern bestenfalls „early“, „continuously“ und „actively“. Ebenso Döring und Bortz (2016, S. 981f.), die die Einbindung als Voraussetzung für eine gute Evaluation sehen und dies *erkenntnistheoretisch*, d.h. dass Evaluationsgegenstände nur dann umfassend betrachtet werden können, wenn auch Beteiligte und Betroffene berücksichtigt werden, *forschungsethisch*, was z.B. die faire Berücksichtigung von Minderheiten oder die transparente Informationsweitergabe beinhaltet, und mit der *besseren Nutzung von Ergebnissen* sowie der *Emanzipation* und dem *Empowerment* von Akteur*innen begründen.

Zwangsläufig können Evaluierende es jedoch nie allen Beteiligten ‚recht machen‘ bzw. allen Interessen nachkommen. Da Evaluationen auch Probleme, Defizite und Fehlentwicklungen aufdecken sollen, werden nie alle Stakeholdergruppen von den Evaluationsergebnissen begeistert sein. Aufgrund dieser Situation sehen Evaluierende sich und/oder ihre Arbeit häufig *Kritik* ausgesetzt, bspw. dass die Ergebnisse keine neuen Erkenntnisse liefern, methodisch unzureichend sind oder in der Durchführung mangelhaft gearbeitet wurde (Stockmann, 2006; Brandt, 2007).¹²⁷ Um möglichst viele Akteur*innen zu erreichen, können *partizipative oder stakeholderorientierte Evaluationsansätze* (s. Kap. 3.1.2) als Vorgehensweise gewählt werden (vgl. Brandt, 2007, S. 271ff.; Döring & Bortz, 2016, S. 1012ff.; Stockmann, 2006, S. 174ff.).

3.3.3 Zur Rolle von Evaluierenden

Personen, die Evaluationen durchführen oder mit diesen beauftragt werden, sind mit zahlreichen Handlungsanforderungen konfrontiert und müssen dadurch unterschiedliche Rollen einnehmen: Sie sind *Wissenschaftler*innen*, *Beratende* (Prozess- und Expertenberatende), *Lehrpersonen* (bspw. zum Anstoß von Lernprozessen), *Richter*innen* (aufgrund der Bewertung i.R. von Evaluationen), *Prozessbegleitende* und *Moderierende*

¹²⁷ Stockmann (2006, S. 268ff.) führt weitere typische Kritikpunkte auf und erläutert sie im Detail. Zudem werden im Anschluss Empfehlungen ausgesprochen, wie Evaluierende sich hiervor schützen können.

(Gutknecht-Gmeiner, 2014; Scriven, 2008). Die zahlreichen Rollen bedingen *Spannungsfelder*, können zu Konflikten führen und fordern umfangreiche *Kompetenzen* seitens der Evaluierenden: Neben allgemeinen evaluationsspezifischen Kompetenzen (z.B. Planung, Durchführung von Evaluationen), methodischen Kompetenzen (z.B. Methoden der empirischen Sozialforschung, der Bewertung, des Projektmanagements), kontextbezogene Kompetenzen (z.B. fachliches Wissen im jeweiligen Praxisfeld) sind vor allem auch *Sozial- und Selbstkompetenzen* gefragt, um u.a. mit den Anforderungen der unterschiedlichen Interessengruppen umgehen zu können (Brandt, 2007).¹²⁸ Auch die unterschiedlichen Werdegänge von Evaluierenden und ihre Verwurzelung in den verschiedensten Wissenschaftsdisziplinen, was die Arbeitsweise letztendlich stark beeinflusst, bestimmen die Diversität dieses Arbeitsfeldes (vgl. Rossi et al., 2004, S. 27.f, S. 372).¹²⁹ In ihrer Profession sind sie zahlreichen Spannungsfeldern ausgesetzt; eine angemessene Position zwischen ihrer Rolle „[...] als *Dienstleistende* für den Auftraggeber einerseits und dem *Anspruch wissenschaftlicher Unabhängigkeit* andererseits [...]“ zu finden, ist nur eines von vielen (Döring & Bortz, 2016, S. 976).¹³⁰

Anders als ältere Positionen gehen neuere Veröffentlichungen nicht von abgrenzbaren und stabilen Rollenverhältnissen aus, sondern von der *Wandelbarkeit oder Gleichzeitigkeit von Rollen bzw. Handlungsanforderungen*: „Dieser Zugang beinhaltet nicht nur eine zeitliche Veränderbarkeit, sondern auch eine gleichzeitige Bezugnahme auf multiple, ggf. ambivalente Handlungsanforderungen“ (Heinze et al., 2019, S. 279). In der Fachdiskussion existiert in diesem Zusammenhang eine „Leerstelle“, was die theoretische und empirische Konzeptualisierung angeht. Heinze et al. (2019, S. 279) sind der Ansicht, dass sich die multiplen Handlungsanforderungen an Evaluierende „[...] nicht durch eine frühzeitige Festlegung und die konsequente Ausübung von spezifischen, feststehenden Rollen begegnen lässt“, sondern sehen *drei Strategien* als erforderlich und hilfreich: (1) „die Schaffung und Gestaltung geeigneter Aushandlungsorte, (2) die Etablierung von Formaten zur (Auf-)Klärung von Erwartungen und Ansprüchen der Handlungspartner

¹²⁸ S. hierzu eine detaillierte Auflistung der Kompetenzen bei Brandt (2007, S. 189) und Anforderungsprofile der DeGEval (2008). Die DeGEval (2008, S. 9ff.) ordnet hierbei Evaluationsstandards den unterschiedlichen Kompetenzbereichen von Evaluierenden zu, d.h. jeder Kompetenzbereich deckt unterschiedliche Standards ab. Zu Anforderungen an Evaluierende allgemeiner s. Döring und Bortz (2016, S. 992).

¹²⁹ Rossi et al. (2004, S. 393ff.) beschreiben „The Profession of Evaluation“ in ihrer Diversität im Detail.

¹³⁰ Hervorhebungen durch die Autorin.

und (3) eine situativ angepasste, wandelbare und zum Teil (personell getrennte) gleichzeitige Übernahme von (ggf. ambivalenten) Rollen von Evaluierenden“.

Da Rollen nicht beliebig situativ angepasst werden können, sollten sie mit beteiligten Stakeholdern in einem für alle nachvollziehbaren Rahmen abgesteckt und definiert werden, bspw. im Einklang mit den DeGEval-Standards. Grundsätzliche Evaluationszwecke und -funktionen und damit verbunden Handlungsanforderungen sollten (möglichst vorab) beschrieben werden (Heinze et al., 2019)¹³¹: „Erst auf dieser Basis können die verschiedenen Erwartungen und Ansprüche an Evaluierende geklärt und eine Akzeptanz der Evaluation bei den beteiligten Akteur*innen erreicht werden, wenngleich auch Evaluation dabei selbst zwangsläufig zum Gegenstand von Aushandlungen und ggf. Adaptationen wird“ (ebd., S. 280).

Der Begriff *Unabhängigkeit*, der mit dem Verständnis eines unabhängigen, objektiven Forschenden einhergeht, wird in den DeGEval Standards nur unter dem Nützlichkeitsstandard im Zusammenhang mit der „Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin“ (DeGEval, 2016, S. 36) und unter dem Fairnesstandard zur „unparteiischen Durchführung und Berichtserstattung“ (ebd., S. 42f.) erwähnt. Letzterer gibt den Hinweis, dass „[...] ein zu enges Verhältnis zu Auftraggebenden, aber auch beispielsweise zu den für den Evaluationsgegenstand verantwortlichen Personen zu vermeiden [ist]“ (ebd., S. 43). Doch durch die Tatsache, dass Auftraggebende zeitgleich Geldgebende von Evaluationen sind, entstehen vielfältige *Einfluss- und Manipulationsmöglichkeiten*, „[...] die im deutschsprachigen Raum nicht selten vorkommen“ (Stockmann, Meyer & Schenke, 2011, S. 39). In einer Umfrage geben bspw. eine Vielzahl der befrag-

¹³¹ Vgl. bspw. Standard *N2 Nützlichkeit* (DeGEval, 2016, S.5) und Standard *B4 Evaluationsvereinbarung zur Planung und Durchführung* der Schweizer Evaluationsgesellschaft SEVAL (2016, S. 5): „Auftraggebende und Evaluierende verständigen sich möglichst früh und verbindlich über die wesentlichen Elemente einer Evaluation und halten diese schriftlich fest, insbesondere Zweck, Gegenstand, Fragestellungen und Methodik der Evaluation, Pflichten und Rechte der Beteiligten, für die Evaluation verfügbare Ressourcen, Fristen für die Erreichung bestimmter Teilergebnisse und die Berichterstattung sowie Offenlegung und Kommunikation der Evaluation und ihrer Ergebnisse.“

ten Evaluierenden (75%) an, dass Auftraggebende verlangten, Ergebnisse positiver darzustellen, dem mehr als die Hälfte (54%) auch nachgingen (ebd., S. 54, 58).¹³² Ausschlaggebende Faktoren für die Beeinflussung können *personenbezogene Merkmale* (bspw. persönliche Betroffenheit und Wahrnehmung; Anzahl durch die Person durchgeführte Evaluationen), *kontextspezifische Merkmale* bzgl. der Evaluierenden (bspw. finanzielle Abhängigkeit von Aufträgen) und *sektorenspezifische Merkmale* (bspw. Sektoren EZ, Hochschule, Berufliche Bildung etc.) sein (ebd., S. 61f.).

Generell sind Evaluierende dazu angehalten, regelmäßig *ihr persönliches Verhalten, ihre Einstellungen und ihre Rolle zu reflektieren*. U.a. *ethnozentrische Sichtweisen*, was bspw. den unreflektierten und unzureichend adaptierten Einsatz von Methoden, Theorien und Instrumenten in einem fremden, kulturellen Kontext meint, oder ‚*going-native*‘, was die Tendenz beschreibt, die Sicht einzelner Stakeholder zu übernehmen, können Störeffekte i.R. der Evaluationsprozesse hervorrufen (Brandt, 2007; Stockmann, 2006).

¹³² In der Studie wurden 260 deutsche Evaluierende mittels einer Online-Befragung zu ihren moralischen Vorstellungen von Unabhängigkeit bzgl. ihrer Profession, zur Häufigkeit, Intensität, subjektiven Wahrnehmung und zur Häufigkeit von negativen Beeinflussungen befragt (Stockmann et al, 2011, S. 44ff.).

3.4 Untersuchungsperspektiven

Über die verschiedenen Zugänge aus der Evaluationsforschung werden die *Evaluationsaktivitäten* im jeweiligen Kontext (EZ, SfE, Praxisbeispiel) als erster Untersuchungsschwerpunkt der Arbeit betrachtet. Die eingeführten *definitiven Betrachtungen* geben als Richtschnur Auskunft darüber, wie die Termini *Evaluation, Forschung, Monitoring, wissenschaftliche Begleitung* sowie andere genutzte Begriffe im jeweiligen Kontext zusammenspielen, verstanden und verwendet werden.

Um die Vielzahl existenter Konzepte, Modelle, Theorien und Ansätze von Evaluation fassen und systematisieren zu können, wird ein *Klassifizierungssystem* ausgewählt, das auf bereits vorhandene, bekannte Systeme aufbaut. Das System nach Rabie (2014) verfolgt einen *pragmatischen Ansatz*: Die erste Kategorie hilft bei der Abgrenzung dessen, was evaluiert wird, also entweder *umfassende Evaluanden* (bspw. Strategien, Systeme, Sektoren, Maßnahmen in ihrer Gesamtheit), die mittels unterschiedlicher Formen von *Reviews* (Evaluationssynthese, Systematic Review, Meta-Evaluation) erfasst werden, oder *Teilbereiche von Maßnahmen*. Die *Evaluation von Maßnahmen*, womit Projekte und Programme in dieser Arbeit bezeichnet werden, macht einen großen Bereich in der Evaluationsforschung aus, so auch in den zu betrachtenden Kontexten. Sie kann sich im Zeitpunkt und ihrer Zielsetzung unterscheiden: *Ex-ante Evaluationen* als Machbarkeitsstudie oder Grundlagenuntersuchung vor einem Projekt, *on-going Evaluationen* zur Prozessevaluation, *Schlussevaluationen* am Projektende zur direkten Zielerreichung oder *ex-post Evaluationen* nach Abschluss einer Maßnahme zur Evaluation von Wirkungen und Nachhaltigkeit. Die Betrachtung von Maßnahmen in solch einer *Verlaufsperspektive* hat den Vorteil, dass (1) auch die Nachförderphase (ex-post) betrachtet wird, in der sich die *Wirkungen und die Nachhaltigkeit* einer Maßnahme herauskristallisieren und (2) die *kausale Verkettung der unterschiedlichen Phasen* hervorgehoben wird, was verdeutlicht, dass die Planung, geschaffene Strukturen und Implementierungsaspekte das Fundament für langfristige Wirkungen und die Nachhaltigkeit der Maßnahme bilden. Mit dem Ziel, möglichst genaue Ursache-Wirkungszusammenhänge darzustellen, sind *Wirkungsevaluationen* mit zahlreichen methodischen und praktischen Herausforderungen (*Störfaktoren*) verbunden und komplexe Designs notwendig. Nachhaltigkeitsevaluatio-

nen, die über die Dauerhaftigkeit von Wirkungen hinausgehen, werden bislang vernachlässigt. Sowohl die Wirkungsmöglichkeiten als auch die Nachhaltigkeit einer Maßnahme sind durch *interne Faktoren* einer Durchführungsorganisation (bspw. Mitarbeiter*innen, Organisationsstruktur etc.) und *externe (Umwelt-)Faktoren* (bspw. andere Systeme, die das Programm umgeben) beeinflusst.

Die Ansätze der zweiten Kategorie helfen bei der *Klärung des Zwecks (warum)* der Evaluation. Dabei zielen *theoriegeleitete Ansätze* darauf ab, das *Wissen über den Evaluationsgegenstand zu erweitern*, bspw. mittels eines *Logischen Modells/Logical Frameworks*, das Aussagen darüber treffen kann, ob vorformulierte Zielindikatoren einer Maßnahme zu dessen Ressourcen (*Inputs*), Leistungen (*Outputs*), Wirkungen auf der Zielgruppenebene (*Outcomes*) oder Wirkungen auf gesellschaftlicher Ebene (*Impacts*) erreicht werden. *Partizipative Ansätze* haben das Ziel, Beteiligte von Maßnahmen i.R. von Evaluationen aktiv miteinzubeziehen, ihre Evaluationskapazitäten befähigen und ein gemeinsames Verständnis zu Sachverhalten zu schaffen. Die dritte und letzte Kategorie beantwortet die Frage, *wie* etwas evaluiert werden soll. *Quantitative Designs* können experimentell oder quasi-experimentell sein. Bei ihnen wird die Zielgruppe einer Maßnahme („mit Intervention“) mit Kontrollgruppen („ohne Intervention“) verglichen. Der Unterschied an Veränderung zwischen den Gruppen kann dann letztendlich als direkte Wirkung der Maßnahme verstanden werden. Bei *RCTs* geschieht die Zuordnung zur Gruppe nach Zufallsprinzip (Randomisierung). *Qualitative Designs* zielen darauf ab, vielmehr den Sinn, den Kontext oder (umgebende, beeinflussende) Prozesse einer Maßnahme zu verstehen, wobei qualitative Befragungsformen (z.B. Interviews, Fokusgruppendifkussionen, Beobachtungen) genutzt werden und bspw. i.R. der *PAR* beteiligte Stakeholder aktiv in Prozesse eingebunden werden. Bei *Mixed-Methods-Designs* werden quantitative mit qualitativen Methoden kombiniert, um Stärken und Schwächen der jeweiligen Methoden auszugleichen und Ergebnisse zu triangulieren.

Von Interesse ist zudem, welche *Funktionen* oder welchen *Nutzen* Evaluationen in den jeweiligen Kontexten verfolgen: Untersucht werden will, ob Evaluationen eher (1) instrumentell eingesetzt werden, d.h. zur *Steuerung* und *Kontrolle* von Maßnahmen, wozu auch die Funktion der Optimierung und Entscheidungsfindung gehören; oder ob sie (2)

eher aufklärend-konzeptuelle Funktionen haben, d.h. die *Erkenntnis* und *Entwicklung* von Maßnahmen im Vordergrund stehen, was auch eine Aufklärungs-, Lern- und Dialogfunktion implizieren kann; oder ob Evaluationen (3) eher taktisch zur *Legitimation* bzw. Rechenschaftspflicht (*accountability*) einer Maßnahme genutzt werden.

Bei der Vielzahl an Evaluationsaktivitäten darf die *Qualität* nicht außer Acht gelassen werden, die *Evaluationsstandards* gewährleisten. Sie beinhalten Kriterien zur *Nützlichkeit*, *Durchführbarkeit*, *Fairness* und *Genauigkeit* von Evaluationen und bedenken die *Rolle von Evaluierenden und anderen Beteiligten* (bspw. im Hinblick auf unparteiische Durchführung und Berichterstattung). *Meta-Evaluationen* als Teil der Evaluationsstandards der DeGEval (2016) haben die Funktion, Evaluationen zu evaluieren und lassen Aussagen zur Qualität der Evaluationen zu.

Dadurch, dass die Evaluationsstandards auch die Evaluierenden und die am Evaluationsprozess Beteiligten ins Blickfeld nehmen, stellen sie eine Art ‚Brücke‘ zum zweiten Untersuchungsschwerpunkt der Arbeit dar, den *Akteur*innen bzw. Stakeholdern*, mit dem Fokus auf den *Evaluierenden*. Generell wird zwischen *interner* und *externer* Evaluation unterschieden, was die von den Organisationen selbst vorgenommenen Evaluationen (*Selbstevaluation*) von jenen, die durch extern beauftragte Evaluierende oder Gutachter*innen durchgeführt werden (*Fremdevaluation*), abgrenzt. Beide Formate bergen *Vor- und Nachteile*, die durch die Kombination der beiden Vorgehensweisen ausgeglichen werden können. Am Evaluationsprozess selbst sind i.d.R. zahlreiche Akteur*innen eingebunden: Neben den Evaluierenden, sind Auftraggebende, Geldgebende, Projekt- oder Programmmanager*innen, Zielgruppen der Maßnahme und lokale Partnerorganisationen (bspw. NGOs, Schulen) mal in passiveren, mal in aktiveren Rollen am Evaluationsgeschehen beteiligt. Unterschiedliche Interessen und Erwartungen der Gruppen führen zwangsläufig zu Spannungen und Konflikten, die Evaluationsprozesse möglicherweise auch behindern können. Die von Beginn an *aktive Einbindung* der Stakeholder (*Partizipation*) erhöht die Transparenz und Akzeptanz von Prozessen und ist neben der Empfehlung von Evaluationsexpert*innen auch in den Evaluationsstandards verankert (*Nützlichkeitsstandards*). Evaluierende sind in diesem sozialen Konstrukt damit konfron-

tiert, sich i.R. von *sich wandelnden und gleichzeitigen Rollen und Handlungsanforderungen* zurecht finden zu müssen. Hierfür sind v.a. *Sozial- und Selbstkompetenzen* gefragt, um mit den unterschiedlichen Interessengruppen zurecht zu kommen und sich darüber hinaus in anderen Spannungsfeldern (*Dienstleiter*in* vs. *Wissenschaftler*in*) positionieren zu können. Auch in diesem Zusammenhang können Evaluationsstandards Orientierung bieten, u.a. um *Rollen und Zuständigkeiten* klar zu formulieren, die *Unabhängigkeit* als Evaluierende*r zu wahren oder mögliche *Einfluss- und Manipulationsmöglichkeiten* seitens der Auftraggebenden zu umgehen. Um Spannungen mit Beteiligten und Störeffekte bei Evaluationen zu vermeiden, ist eine ausgeprägte *Reflexionsfähigkeit* der Evaluierenden bzgl. des eigenen Verhaltens, Einstellungen und von Rollen gefragt. Folgende Abbildung fasst alle Untersuchungsperspektiven zusammen, wobei gestrichelte Pfeile für Beeinflussungsprozesse stehen:

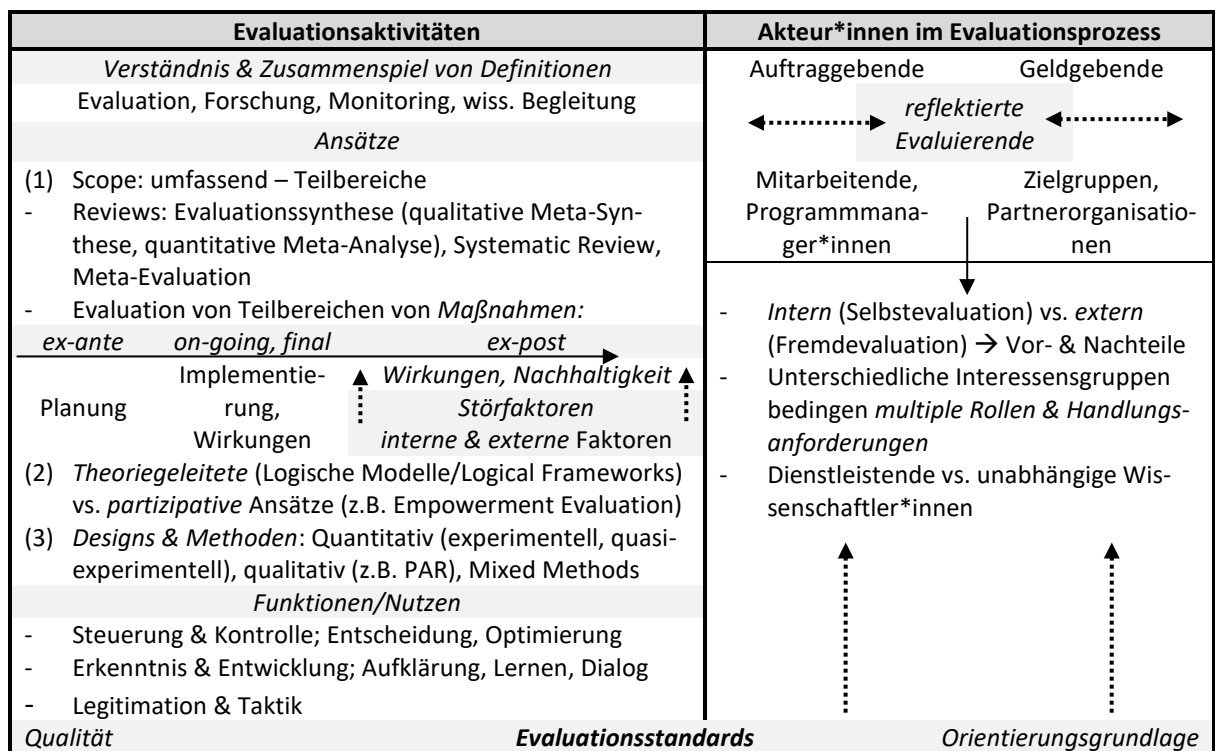


Abb. 10: Theoretisch-konzeptionelle Rahmung der Arbeit (eigene Darstellung)

3.5 Methodische Vorgehensweise: Literatur- und Dokumentenrecherche

Zu den genutzten Quellen für diese Arbeit in deutscher und englischer Sprache zählen sowohl Primär- als auch Sekundärliteratur. Konkret sind dies Sammelwerke, Monographien und Fachzeitschriften sowie Beiträge, graue Literatur/Berichte/Reports (z.B. Resolutionen und Beschlüsse der Vereinten Nationen und der Bundesregierung, Internetdokumente in gedruckter Form), Hochschulschriften, Internetdokumente (Webseiten) und Protokolle. Durch die wissenschaftliche Begleitung stehen darüber hinaus vereinzelt auch interne (Evaluations-)Berichte zum SV SfE der GIZ zur Verfügung, insofern dies vertraglich vereinbart wurde.

Bei der Literatur- und Dokumentenrecherche für die sportbezogenen Inhalte dieser Arbeit (*Evaluation und Forschung im Kontext von SfE, Internationale Sportförderung und SfE in Deutschland*) werden in einem ersten Schritt die über die Zentralbibliothek der DSHS zugänglichen internen sowie externen Kataloge und Datenbanken genutzt. Zu den internen oder frei zugänglichen Datenbanken gehören der *Gesamtkatalog der ZBSport* sowie *SURF*, das Sportinformationsportal des Bundesinstituts für Sportwissenschaft (BiSp), das die Datenbanken *Spolit*, *Spofor*, *Spomedia* und den Fachinformationsführer *Sport* beinhaltet, wobei v.a. *Spolit* die Recherche von sportbezogenen Aufsätzen aus Zeitschriften und Sammelwerken unterstützt. Über *SURF* ist bspw. auch der Zugang zu allen *Sportberichten der Bundesregierung* seit 1970 gewährt. Darüber hinaus wird die *Datenbank Bielefeld Academic Search Engine (BASE)*, eine der weltweit größten Suchmaschinen für wissenschaftliche Web-Dokumente, und vor allem die *Digitale Bibliothek (DigiBib)*, bei der über Fernleihe auf den Zeitschriften- und Bücherbestand aller größeren Bibliotheken in Deutschland und im Ausland Zugriff besteht, für die weiterführende Recherche verwendet, d.h. für die sportunspezifischen Bereiche der *allgemeinen Evaluationsforschung* und der *Evaluation in der EZ*. Zudem kommen die frei zugängliche Suchmaschine *GoogleScholar* für wissenschaftliche Dokumente allgemein und *GoogleBooks* als Sammlung von retrodigitalisierten Beiträgen am Rande zum Einsatz.

Des Weiteren bedingen die zu untersuchenden Bereiche auch unterschiedliche Zugänge bei der Literaturrecherche. So wird für den Bereich der *EZ-Evaluation* die *OECD-i-Library* durchsucht, zu der über die Universität zu Köln ein Zugang besteht. Des Weiteren wird

das *Publikationsportal DEval und der GIZ* genutzt. Die GIZ hat zudem eine separate Datenbank für ihre durchgeführten Evaluationen veröffentlicht, d.h. eine frei zugängliche Excel Tabelle inklusive Filterfunktionen, die Evaluationsberichte des Unternehmens seit ca. 2007 listet.

Einschlägige Fachzeitschriften aus dem Bereich der *Evaluationsforschung* werden auf Beiträge zur Entwicklungszusammenarbeit bzw. Entwicklungspolitik durchsucht, aber auch zu Stichworten mit allgemeiner Relevanz, wie ‚Wirkung*/impact‘ und ‚Nachhaltigkeit/sustainability‘. So bspw. die *Zeitschrift für Evaluation*, die seit 2002 „[...] wissenschaftliche Originalbeiträge zu Theorie, Methoden und Praxis der Evaluation als auch praxisorientierte Erfahrungsberichte [...]“, veröffentlicht. „Darüber hinaus gibt es Debattebeiträge, Tagungsberichte, Rezensionen aktueller (Buch-)Veröffentlichungen, Präsentationen von Evaluationsinstitutionen“ sowie eine Rubrik mit Informationen über aktuelle Aktivitäten der Gesellschaft für Evaluation (DeGEval) (DeGEval, 2021, 9. Juni).

Für den internationalen Untersuchungsbereich von SfE bietet das frei zugängliche *Journal of Sport for Development*, das seit 2013 SfE-spezifische Beiträge zu unterschiedlichen Themenschwerpunkten sowie Sonderausgaben online veröffentlicht, einschlägige Recherchemöglichkeiten. Hier werden alle Ausgaben in Augenschein genommen. Darüber hinaus werden *Fachzeitschriften anderer sportwissenschaftlicher Disziplinen*, i.R. derer eine Mehrheit von SfE-bezogenen Artikeln publiziert werden (bspw. *Sport in Society*, *Sport Management Review*, *International Journal of Sport Policy and Politics*), umfangreich durchsucht.

Alle benannten Zugänge werden für jeden Untersuchungsbereich durch *Suchbegriffe und deren Kombinationen* strukturiert. Für den Bereich der *EZ-Evaluation* sind dies die Suchbegriffe *evaluation, research, Forschung, Methode(n)/method(s)*, z.T. in Kombination mit *Entwicklung*(Entwicklungszusammenarbeit, -hilfe), development, cooperation/aid*. Für die Evaluation und Forschung im Kontext von SfE werden die Stichworte *evaluation, research, Forschung, Methode(n)/method(s)* z.T. in Kombination mit oder alleinig die Begriffe *sport + development/Entwicklung, „Sport for Development“, „Sport für Entwicklung“, Sportförderung* verwendet.

Die durch die Suchkriterien ermittelten Beiträge werden anhand der theoretisch-konzeptionellen Untersuchungsperspektiven inhaltlich analysiert. Zitationen jener Beiträge, die inhaltlich für die verschiedenen Kontexte als besonders relevant einzustufen sind, werden für weitere Recherchen genutzt. Folgende Tabelle fasst die genutzten Zugänge dieser Arbeit zusammen:

Tab. 5: Bibliographische Zugänge (eigene Darstellung)

Bibliographische Zugänge	Untersuchungsbereiche			
	Evaluationsfor- schung allg.	Evaluation in der EZ	Evaluation, Forschung im Kontext SfE	Sportförderung SfE in Deutsch- land
Kataloge und Datenbanken				
Online-Katalog „ZB Sport“ der DSHS https://aleph.zbsport.de/F?func=find-b&find_code=WRO&local_base=zbs01&CON_LNG=ger&request=&Button=Suche	X	X	X	X
SURF des BiSp, v.a. Datenbank Spolit: https://www.bisp-surf.de/Search/Results?filter%5B%5D=format%3A%22Literatur			X	X
Digitale Bibliothek (DigiBib): https://dshs-koeln.digibib.net/	X	X	X	X
Bielefeld Academic Search Engine (BASE): https://www.base-search.net/	X	X	X	X
Google Scholar: https://scholar.google.com/ Google Books: https://books.google.de/	X	X	X	X
OECD iLibrary: https://www.oecd-ilibrary.org/		X		
GIZ Publikations-Portal/Datenbank: https://publikationen.giz.de/esearcha/browse.tt.html		X		
GIZ Datenbank Evaluierungen mit Filterfunktionen: https://www.giz.de/de/downloads/GIZ-Datenbank%20begonnene%20und%20abgeschlossene%20Evaluierungen.xlsx		X		
DEval Publikationen: https://www.deval.org/de/publikationen?tx_lfpublications_list%5Boffset%5D=0&cHash=52aa7e3c7d05e8796224b253fe6b315c		X		
Einschlägige Zeitschriften				
European Journal for Sport and Society (ISSN 1613-8171)			X	
Evaluation (ISSN 1356-3890)	X	X		
Evaluation Practice (ISSN 0886-1633)	X	X		
International Journal of Sport Policy and Politics (ISSN 1940-6940)			X	
International Review for the Sociology of Sport (ISSN 1012-6902)			X	
Journal of Sport and Social Issues (ISSN 0193-7235)			X	
Journal of Sport for Development (open access): https://jsfd.org/			X	
Journal of Sport Management (ISSN 0888-4773)			X	
Sport, Education and Society (ISSN 1357-3322)			X	

3 Rahmung

Sport Management Review (ISSN 1441-3523)			X	
Sport in Society (ISSN 1743-0445)			X	
Sociology of Sport Journal (ISSN: 0741-1235)			X	
Qualitative Research in Sport, Exercise and Health (ISSN 2159-676X)			X	
Zeitschrift für Evaluation (ZfEv) (ISSN 1619-5515)	X	X		
Sonstige Quellen				
Internetseite S4D Resource Toolkit https://www.sport-for-development.com/home				X
Interne (Evaluations-)Berichte				X

4. Betrachtung des EZ-Kontexts: Entwicklung von Evaluation in Phasen

Im folgenden Kapitel wird die *Evaluationspraxis in der EZ* chronologisch beleuchtet. Hierbei werden sowohl internationale als auch nationale Meilensteine in der EZ, Standardisierungsprozesse (bzgl. Terminologien etc.), wegweisende Studien¹³³, thematische Schwerpunkte und Ansätze (u.a. Designs, Methoden), sowie Institutionalisierungs- und Professionalisierungsprozesse der Epochen beschrieben. In der chronologischen Betrachtung wird sowohl die *EZ als Untersuchungsgegenstand* als auch die *Forschung über die Evaluation in der EZ* einbezogen, da sich diese Prozesse im Zeitverlauf wechselseitig beeinflusst haben und teilweise eng miteinander verwoben sind. Durch die Darstellung wird deutlich, wie sich die Evaluationsforschung in der EZ im Laufe der Zeit bzgl. der Untersuchungskomponenten (Definitionen, Ansätze, Funktionen, Qualitätsaspekte, Evaluation von Wirkungen und Nachhaltigkeit) entwickelt hat und welche Akteur*innen (international, national; Evaluierende intern/extern) diese Entwicklungen präg(t)en. Das Kapitel bietet weder einen vollständigen Überblick zur Entwicklung der Evaluationsforschung allgemein¹³⁴, noch deckt es umfangreich Entwicklungen im Politik- und Arbeitsfeld der EZ ab.¹³⁵ Meilensteine in der EZ-Geschichte sowie Diskurse der Disziplin werden lediglich dann aufgegriffen, wenn sie einen direkten Zusammenhang zur bzw. einen direkten Einfluss auf die Evaluationspraxis hatten. Auf eine differenzierte Darstellung zu Entwicklungen in der Entwicklungsforschung wird ebenfalls verzichtet.¹³⁶ Da sich die EZ-

¹³³ Der Begriff *Studien* wird in den weiteren Ausführungen als Sammelbegriff jeglicher Form von Evaluationen und Untersuchungen genutzt und/oder wenn diese nicht eindeutig den gewählten Klassifikationen (s. Kap. 3) zugeordnet werden können.

¹³⁴ Einen Überblick zu Entwicklungen und Entstehung der *Evaluationsforschung* in den USA, Europa und Deutschland bieten die Beiträge in Stockmann (2000a); Stockmann und Meyer (2017a, 67ff.) und Wollmann (2017). Zur Entwicklung und dem Stand der Evaluationsforschung in weiteren Ländern s. im *International Atlas of Evaluation* (Furubo, Rist & Sandahl, 2002) und aktuell zu Ländern Europas Stockmann und Meyer (2020). Furubo (2017) beschäftigt sich mit der Entwicklung von Evaluation allgemein und geht dabei von einem inkrementellen Wandel aus. Chronologisch gibt der Autor (ebd.) auch einen Einblick in die Ursprünge von Evaluation, die zeitlich vor der in dieser Arbeit abgehandelten Chronologie datiert sind.

¹³⁵ Zu *Entwicklungen in der Entwicklungspolitik und EZ* s. u.a. Stockmann, Menzel und Nuscheler (2016), Sasaki (2006), Nowak (2014), Neubert (2016), Eckert (2015) sowie Darnell (2012, S. 40ff.) aus einer kritischeren Perspektive; Klingebiel (2017) und Beiträge in Debiel (2018) vor dem Hintergrund der Agenda 2030. Klingebiel (2013) gibt eine allgemeine *Einführung in die EZ*, inklusive Definitionen, Ansätze, Details zur Wirksamkeitsdebatte und Herausforderungen.

¹³⁶ Für einen allgemeinen *Überblick zur Entwicklungsforschung* s. Currie-Adler, Kanbur, Malone und Medhora (2014), Fischer, Hauck und Boatcă (2016), Rodrick und Rosenzweig (2009). Klingebiel (2017) betrachtet Veränderungen in der Entwicklungsforschung vor dem Hintergrund der Agenda 2030.

Evaluation v.a. in der Gegenwart stark ausdifferenziert hat und dadurch schwer zu überblicken ist, erhebt diese Darstellung keinen Anspruch auf Vollständigkeit. V.a. internationale Prozesse können nicht umfangreich abgebildet werden, der Fokus liegt auf den Entwicklungen in der dt. EZ.

Im historischen Verlauf identifizieren Autor*innen unterschiedliche Phasen bzgl. der Etablierung von Evaluation in der EZ: Cracknell (2002, S. 39f.) unterscheidet in (1) *Early Developments* (1960-1970er Jahre), (2) *Explosion of Interest* (1979-1984), (3) *Coming of Age* (1985-1988) und (4) *Aid Evaluation at the Crossroads* (ab 1988). Wollmann (2017) sowie Stockmann und Meyer (2017a) unterscheiden hinsichtlich der Entwicklung von Evaluation und Evaluationsforschung in Deutschland verschiedene Phasen, die sich an den unterschiedlichen Funktionen und Aufgaben von Evaluation orientieren. Auch wenn Wollmann (2017), Stockmann und Meyer (2017a) Entwicklungen von Evaluation im Allgemeinen betrachten, stehen diese in direktem Zusammenhang mit den Entwicklungen der Evaluation in der EZ als einem Teilbereich der Entwicklungspolitik (vgl. Neubert, 2016, S. 359; Stockmann, Menzel & Nuscheler, 2016, S. 8f, S. 425ff). Kjekshus (1991, S. 60) sieht die Entwicklung von Evaluation eng mit den Entwicklungen in der EZ verwoben: „In Europe, its [evaluation’s] development has [...] been connected almost exclusively with development aid management and administration and with the official aid community“. In Anlehnung an genannte Autor*innen wird im Folgenden die Entwicklung von Evaluation in der EZ chronologisch dargestellt.

4.1 Phase I: Die Anfänge (1960er und 1970er)

Nach Stockmann und Meyer (2017a, S. 69) sollte EZ-Evaluation in dieser Aufbruchsphase v.a. „[...] zur *Verbesserung der Implementation und Outputsteigerung* beitragen“. Cailloux (1988, S. 162) spricht von der „ersten Generation“ in der EZ-Evaluation: „It showed a limited number of (individual) donor efforts. Donors were mainly concentrating on project measurement systems, looking at *performance*, using *audit-oriented techniques*, basically *quantified analysis*, and very few recipients were involved at all.“¹³⁷

Die Ursprünge von Evaluation als „tool in public administration“ (Berlage & Stokke, 1992, S. 2) liegen in den Vereinigten Staaten (Cracknell, 2002; Kjekshus, 1991). In den 1960er Jahren wurden seitens der amerikanischen Entwicklungsagentur *USAID (United States Agency for International Development)* und in wenigen größeren UN-Organisationen erste Versuche unternommen, Evaluation zu einem festen Bestandteil des Projekt- und Programmmanagements zu etablieren (Hageboeck, 1984), allerdings hauptsächlich lediglich in der Endphase eines Projekt-/Programmzyklus (Cracknell, 2002). Durch die Entwicklung des theoriegeleiteten *Logical Frameworks* sollte das Programmdesign unterstützt und Erfolgsfaktoren und Projektergebnisse messbar gemacht werden. Im weiteren Verlauf wurden solche Ansätze von europäischen Akteur*innen übernommen (Berlage & Stokke, 1992), laut Cracknell (2002) dauerte dies allerdings eine ganze Dekade.¹³⁸ Generell fand die Entwicklung der EZ-Evaluation in europäischen Ländern äußerst isoliert statt: „[...] the growth and coming of age of the evaluation discipline in European countries' aid organizations seem to have taken place [...] in response to unique national needs and circumstances, administrative traditions and political pressures“ (Kjekshus, 1991, S. 60).

Durch die beginnende Nutzung des *Logical Frameworks*, fokussierten sich Studien in dieser Phase hauptsächlich auf *Kosten-Nutzen-Analysen* zur Bewertung (z.B. McKean, 1966; Dorfman, 1963; Mishan, 1971) und damit auf *Steuerung* und *Kontrolle* von Maßnahmen. Die Betrachtung und Bewertung von EZ-Maßnahmen aus ökonomischer Sicht stand im

¹³⁷ Hervorhebungen durch die Autorin.

¹³⁸ In Cracknell (2002, S. 101ff.) kann nachgelesen werden, ab wann welche Akteur*innen das *Logical Framework System* im Laufe der Zeit übernommen bzw. adaptiert haben.

Mittelpunkt, was das Manual „A Guide to the Economic Appraisal of Projects in Developing Countries“ (ODA, 1975) zum weltweiten ‚Verkaufsschlager‘ machte. Des Weiteren stieg das Interesse an Untersuchungsdesigns, die das „with-and-without“, also das *Kontrafaktische*, mit einbezogen als lediglich „before-and-after“-Vergleiche zu implizieren (Cracknell, 1988). Auch wenn die Ergebnisse von Hirschmans *Wirkungsevaluation* mit dem Titel „Development Projects Observed“ (1967) eher enttäuschend ausfielen, erlangte die Studie große Aufmerksamkeit und wird von Cracknell (2002, p. 40) als „the true preliminary trial of evaluating aid projects“ bezeichnet.¹³⁹

Bzgl. der Evaluationsforschung in Deutschland nahm das BMZ im Vergleich zu Repräsentant*innen anderer politischer Sektoren eine Vorreiterrolle ein: Zwischen der *Kreditanstalt für Wiederaufbau (KfW)* und dem BMZ wurden bereits 1966 sog. *Erfolgskontrollen* aller EZ-Maßnahmen vereinbart (vgl. Terberger, 2011, S. 223). Ab 1971 wurde im BMZ ein *Inspektionsreferat* etabliert, das später als *Zentrale Erfolgskontrolle* die Evaluations-einheit im Ministerium darstellte (Wollmann, 2017). Gemäß der Namensgebung war diese Einheit vor allem eine *Kontrollinstanz*, „[...] die neben einer stark ökonomisch fokussierten Wirkungsanalyse an prominenter Stelle auch die ordnungsgemäße Durchführung von Maßnahmen zu überprüfen hatte“ (Bliss, 2006, S. 14). Für die frühe Institutionalisierung der Evaluationspraxis in der dt. EZ sieht Lachenmann (1977, S. 68f.) vier Gründe: (1) Die EZ musste sich gegenüber kritischen Entwicklungstheorien *rechtfertigen*; (2) Wirkungen und Nutzen mussten dem deutschen Steuerzahler *nachgewiesen* werden; (3) zur Kontrolle der Maßnahme im Partnerland und (4) die enge Verbindung zur US-amerikanischen EZ. Demnach waren die *Legitimations-* (1, 2) sowie die *Kontrollfunktion* (3) zentrale Treiber für Institutionalisierungsprozesse und damalige Hauptfunktionen von EZ-Evaluationen.

Im Zuge einer internationalen Konferenz der *UNESCO (United Nations Educational, Scientific and Cultural Organization)* im Jahr 1977 wurden erstmalig Expert*innen geladen, um gemeinsam den *Bedarf an systematischerer Evaluation* zu reflektieren und einen Einblick in den aktuellen Stand der Evaluationsforschung zu bekommen (Sasaki, 2006).

¹³⁹ Cracknell (1988; 2002) beschreibt in dieser Phase ausführlicher die Situation in der englischen Geberinstitution *ODA (Overseas Department Administration)* sowie Studien, die in dieser Zeit durchgeführt wurden, bspw. auch erste Grundlagenuntersuchungen in Ländern.

Basierend auf dem Vortrag wurden die Handbücher „Evaluation – A systematic Approach“ (Rossi, Freeman & Wright, 1978)¹⁴⁰ und „Evaluating Social Projects in Developing Countries“ (Freeman, Rossi & Wright, 1979) veröffentlicht, die unter EZ-Akteur*innen bald weit verbreitet waren (vgl. Sasaki, 2006, S. 66).

Bis Ende der 1970er blieb das Thema Evaluation in der EZ allerdings „still in its infancy“ (Cracknell, 1988, S. 75). Neben den Impulsen aus den USA seitens der Weltbank, USAID sowie weniger größerer UN-Organisationen, „[...] did [the OECD] some useful work in bringing evaluators together“ (ebd., S. 75). Hiermit sind *erste Konferenzen und Seminare* zum Thema gemeint¹⁴¹, sowie *wenige erste Leitfäden*. Bspw. werden im Bericht „Evaluating Development Assistance“ von 1972 methodische Probleme beschrieben und Empfehlungen gegeben, wie Evaluation als effektives „management tool“ genutzt werden kann (OECD, 2013, S. 6). Hierbei hat laut Cracknell (2002, S. 40) die europäische Evaluationsgemeinschaft rückblickend einen wichtigen Beitrag geleistet „[...] in getting things moving on the evaluation front in those early days“. Dass Evaluation ein wichtiges Thema auf der Agenda des Staatenbündnisses OECD ist, wurde bereits ein Jahr nach der Gründung im Jahr 1961 deutlich: „Member Governments would be well advised to devote more attention to critical evaluation of the results of their aid operations and exchange experience in this field“ (DAC, 1962, S. 41 zitiert nach OECD, 2013, S. 6). Jedoch dauerte es noch weitere zwanzig Jahre bis zur Gründung der *DAC Expert Group on Aid Evaluation* im Jahr 1982 (ebd.).

¹⁴⁰ Das Buch erschien in zahlreichen weiteren Auflagen und ist heute noch aktuell. Es wird auch i.R. dieser Arbeit genutzt.

¹⁴¹ In Deutschland (1966), den Niederlanden (1970 und 1973) und Dänemark (1975) (OECD, 2013).

4.2 Phase II: Rasante Entwicklungen und erste Krisen (1980er)

Die 1980er Jahre stehen für einen rasanten Aufschwung der EZ-Evaluation und gleichzeitig für erste schwierige Zeiten:¹⁴² Mitte der 1980er steckte die Evaluation in einer *dreifachen Krise*:

„Die Ergebnisse aus Evaluierungen würden nicht ausreichend genutzt (crisis of utilization), ihre Validität und Bedeutsamkeit seien oft zweifelhaft (crisis of verification) und die Werturteile, die in die Forschung Eingang fänden, würden zu wenig reflektiert und expliziert (crisis of accreditation) (Neubert, 2010, S. 141).“

Nach Stockmann und Meyer (2017a, S. 69) wird Evaluation auch in dieser Phase vor allem als „taktisches Instrument“ zur *Legitimierung von politischen Entscheidungen* genutzt. Laut Wollmann (2017, S. 42) kam ihr vorrangig die Aufgabe zu, „[...] Kosten (Input) zu reduzieren und zu minimieren“.

Ab Ende der 1970er stieg die Zahl an Evaluationseinheiten oder -abteilungen in europäischen EZ-Organisationen sprunghaft an (vgl. Stokke, 1991, S. 1; Cracknell, 2002, S. 45), weswegen auch von einem ersten *Evaluations-Boom* gesprochen wird (Sasaki, 2006, S. 63; Stockmann, 2017, S. 36). In den frühen 1980ern hatte nahezu jede Geberinstitution eine eigene Abteilung, die primär oder alleinig für Evaluation verantwortlich war (vgl. Stokke, 1991, S. 8). Grund dafür waren vorangehende finanzielle Kürzungen bei EZ-Programmen, was zu einem gesteigerten Interesse an Evaluationen führte und zur Aufstockung von Evaluationskapazitäten in den EZ-Organisationen (Cracknell, 2002).¹⁴³ Die *Frage nach der Wirksamkeit* oder (*Aid*) *Effectiveness*¹⁴⁴ von EZ wurde zum ersten Mal adressiert: „The question raised by aid managers and politicians was clearly about the impact of aid. [...] The challenge was: could the evaluators give a meaningful answer?“ (Kjekshus, 1991, S. 61). Die Frage „does aid work?“ konnten die Evaluierenden der OECD durch eine erste Studie aufgrund der mangelhaften Datenlage nicht beantworten, die

¹⁴² MacArthur (1995) und Stokke (1991) geben zu den 1980er/Anfang 1990er Jahre einen guten Überblick zur Evaluationspraxis allgemein und der genutzten Ansätze und Methoden in einzelnen Entwicklungsagenturen.

¹⁴³ Bspw. die englische Geberinstitution (*Overseas Development Administration, ODA*) verdreifachte ihre Ausgaben für Evaluationen und verdoppelte das Personal, was zu einem erheblichen Anstieg in der Qualität und Quantität von Entwicklungsberichten führte (Cracknell, 2002).

¹⁴⁴ Die Begriffe *Wirksamkeit* und (*Aid*) *Effectiveness* werden im Folgenden gleichbedeutend verwendet.

u.a. der heterogenen Evaluationspraxis zu dieser Zeit geschuldet war.¹⁴⁵ Basierend auf diesen ersten aufkommenden Wirksamkeitsdebatten und den enttäuschenden Evaluationsergebnissen gründete die OECD 1982 die *DAC (Development Assistance Committee) Expert Group on Aid Evaluation*¹⁴⁶ und gab damit dem Thema Evaluation eine neue Gewichtung – und etablierte zeitgleich ein neues Forum für Debatten, Austausch und methodische Weiterentwicklung:

„The Expert Group is unique in the sense that there is no other international forum where aid evaluation experts meet regularly to exchange experiences and develop methodology. In most agencies aid evaluation represent a very small – and often somewhat isolated – group of people and the Expert Group therefore has an opportunity for the newcomers to the profession to learn [...]“ (OECD, 1993, S. 4)

Das Zusammenführen von Ergebnissen sowie die Garantie eines regelmäßigen Informationsaustausches unter den nationalen EZ-Agenturen waren Hauptaufgaben der Expert*innengruppe (OECD, 1988). Die Ergebnisse der ersten Wirksamkeitsstudie dieser Gruppe wurden zwei Jahre später in Form von *lessons learnt* veröffentlicht. Sie führten zur *Restrukturierung, Reorganisation und Neukonzeptionierung* i.R. der Evaluationseinheiten, was sich u.a. auf die Vorgehensweise und dem Verständnis von Evaluationsfunktionen auswirkte. Bspw. wurde das *Logical Framework* oder Abwandlungen des Modells¹⁴⁷ in zahlreichen EZ-Organisationen eingeführt und es wurde vermehrt zwischen *internen und externen Evaluationen* bzw. *Selbst- und Fremdevaluierung* unterschieden (vgl. Kjekshus, 1991, S. 63). Auch viele europäische Geberinstitutionen richteten ihre Evaluationssysteme in den 1980ern am zielorientierten, theoriegeleiteten *Logical Framework*-Ansatz aus, um eine *Kontrollfunktion* zu gewährleisten (vgl. Berlage & Stokke, 1992, S. 2) und Projekte besser planen und steuern zu können (Cracknell, 2002;

¹⁴⁵ „At the first draft buried the unwelcome news that the aid effectiveness question could not be answered at all, if ever“ (OECD, 1993, S. 10). Im „Report on Aid Effectiveness“ von 1982 wurden die negativen Ergebnisse dafür genutzt, die Handlungsnotwendigkeit bzgl. der Verbesserung von Evaluation hervorzuheben (OECD, 2013).

¹⁴⁶ Im Folgenden als *DAC-Expert*innengruppe* oder *Expert*innengruppe* bezeichnet. Sie wurde noch weitere zwei Mal umbenannt: 1998 in *Working Party on Aid Evaluation* und 2003 in *DAC Network on Development Evaluation* (OECD, 2013). Die Expert*innengruppe setzte sich aus den Leiter*innen der Evaluationseinheiten der nationalen EZ-Agenturen sowie der Weltbank und dem Internationalen Währungsfond als Beobachter*innen zusammen.

¹⁴⁷ Bspw. die *Zielorientierte Projektplanung (ZOPP)* der GTZ, s. Beschreibung im Folgenden.

Sasaki 2006).¹⁴⁸ Das System wurde zu einem unerlässlichen Teil des Projektmanagements in der EZ und als ein großer Fortschritt hinsichtlich der bislang unorganisierten und unsystematischen Methoden bewertet: „For the first time the objectives had to be stated and criteria for measuring progress towards achieving them had to be identified“ (Cracknell, 2002, S. 47). Basierend auf dem Ansatz entwickelte die GTZ die *Zielorientierte Projektplanung (ZOPP)* und vertraute darauf, dass „[...] eine „gute“ Planung zu wirkungsvollen Projekten führen würde“ (Stockmann, 2000b, S. 382). Mit dieser Vorgehensweise war jedoch lediglich eine *starke Input-Output-Orientierung* gegeben, was dazu führte, dass Wirkungen vernachlässigt wurden (ebd.).

Aufkeimende Kritik an der allgemeinen Durchführung und Umsetzung von EZ (vgl. Riddell, 1987; Berlage & Stokke, 1992, S. 2) verlagerte sich auch in den Diskurs um die Evaluationspraxis, wie am Beispiel der ZOPP der GTZ: So „[...] waren Zielsetzungen und Konzeption [...] hauptsächlich an den Interessen und soziokulturellen Verhältnissen der Geber ausgerichtet. Die Partner wurden bei der Einführung [...] nicht gefragt“ (Kohnert, 1998, S. 138).¹⁴⁹ Kritische Wissenschaftler*innen, darunter Chambers (1981), Kamla (1985), Salmen (1987) sowie Guba und Lincoln (1989), waren Pionier*innen im Feld der *Entwicklung und Nutzung partizipativer Ansätze und Methoden* (Cracknell, 2002) wie dem *Participatory Rural Appraisal (PRA)*, „bei denen bewusst auf die praktische Nützlichkeit von Evaluierungen geachtet wurde und Zielgruppen beteiligt wurden“ (Neubert, 2010, S. 141). Die Wichtigkeit dieser „neuen Ideen“, aber auch die aktive Einbindung und Gestaltung von Evaluationsprozessen durch Partner*innen und Zielgruppen, wurde erkannt und war im Verlauf eines weiteren Jahrzehnts Diskussionsgegenstand (Neubert, Mack & Roxin, 2014, S. 62; Cracknell, 2002, S. 47), bevor sie Anfang der 1990er Jahre auf eine breitere Zustimmung in der EZ-Praxis trafen. Die systematische Einbindung der

¹⁴⁸ Hemmer (2007, S. 2ff.) und Binnendijk (2000, S. 16) beschreiben die *Logical Framework-Analyse* als Instrument der EZ im Detail. Cracknell (2002, S. 101ff.) geht auf den Aufbau, die Funktionen und die Nutzung sowie Vorzüge und Nachteile ein. Das *Logical Framework* wurde ob seiner ganzen Vorzüge im Vergleich zur vorherigen Evaluationspraxis einige Zeit wenig kritisch betrachtet. Die Schwächen wurden erst nach und nach identifiziert.

¹⁴⁹ Kohnert (1998) veranschaulicht ausführlicher Stärken und Schwächen der ZOPP.

Partner*innen wird seit den 1980ern zwar von zahlreichen Praktiker*innen gefordert, jedoch bis heute wenig in die Tat umgesetzt (Neubert, 2010).¹⁵⁰

Cailloux (1988) spricht von einer *rasanten Entwicklung und Diversifizierung bzgl. methodischer Vorgehensweisen* in den 1980ern, auch wenn sie kein Schwerpunkt der Arbeit von Geberorganisationen darstellte (vgl. ebd., S. 160ff.). Diskussionen der noch jungen DAC-Expert*innengruppe zur Verbesserung der Evaluationspraxis drehten sich viel um *Evaluationsdesigns und Methoden*. Themen wie die Nutzung des *Logical Frameworks*, die Betrachtung von Projekten in einem weiteren makroökonomischen Zusammenhang, die Schwierigkeiten bei Programmevaluationen, die Nutzung von Kosten-Nutzen-Analysen sowie die Rolle von experimentellen Designs werden die Expert*innengruppe auch über die nächsten Jahrzehnte noch beschäftigen (vgl. OECD, 2013, S. 32). Für Cracknell (2002, S. 48) steht eine Konferenz 1989 für einen Wendepunkt in der EZ-Evaluation: „This conference [...] signaled the beginning of a significant switch in emphasis of evaluation work from the previous emphasis on project management and the logical framework, towards a new emphasis on stakeholder analysis and a willingness to experiment with participatory approaches.“

Basierend auf den Debatten und mit dem Ziel, *Evaluationskapazitäten in den Partnerländern* zu stärken, fand 1987 ein erstes Evaluationsseminar der DAC-Expert*innengruppe mit Repräsentanten aus Entwicklungsländern statt (OECD, 1993, S. 19). Auf Anstoß der Weltbank wurden auch zunehmend M&E-Einheiten in Entwicklungsländern etabliert. Sie waren oft Voraussetzung für eine Finanzierung seitens der Weltbank.¹⁵¹ Die Weltbank nahm eine führende Rolle hinsichtlich der Evaluation von EZ ein, da ihr hierfür mehr Ressourcen als anderen Geberinstitutionen zur Verfügung standen. Interne

¹⁵⁰ Befürworter*innen der partizipativen Ansätze argumentierten, „[...] dass diejenigen Menschen, auf die die Entwicklungsmaßnahmen abzielten, auch ihre Wirkungen am authentischsten beurteilen könnten“ (Neubert et al., 2014, S. 62). Auftraggebende bemängeln an den Verfahren den zeitlichen Aufwand. Anhänger*innen der quantitativen Sozialforschung kritisieren nach wie vor die Verzerrung von Ergebnissen durch die Einbeziehung der Zielgruppen sowie einen unvollständigen Überblick derer hinsichtlich der Wirkungszusammenhänge, „auch wenn diese Einschätzungen bereits häufig in der Praxis widerlegt oder relativiert wurden [...]“ (Neubert, 2010, S. 141).

¹⁵¹ Das Zusammenführen zweier wichtiger und untrennbarer Aufgaben i.R. des Programmmanagements stellte jedoch häufig eine Überforderung dieser Einheiten dar: „masses of data“ wurden zwar gesammelt, allerdings nie ausgewertet (Cracknell, 2000, S. 163). 1986 wurde das Zusammenführen dieser Funktionen in einer Einheit seitens der Weltbank als „a bad mistake“ deklariert (ebd., S. 45).

Betriebsanweisungen und Leitfäden wurden von zahlreichen anderen Institutionen genutzt (ebd.). Generell leisteten Entwicklungsbanken und andere multilaterale Entwicklungsagenturen Pionierarbeit, indem sie früh *eigene Evaluationssysteme* entwickelten (Cracknell, 2002, S. 46).¹⁵² Kleinere Geberorganisationen waren ab Ende der 1980er Jahre zunehmend daran interessiert, mehr Einblicke in die *M&E-Systeme multilateraler Organisationen* zu bekommen, sowohl bzgl. der Wirksamkeit der Organisationen als auch in die internen Evaluationsabläufe und -kapazitäten, anhand derer Entscheidungen getroffen werden. Nur die USA hatten bis dato Wirkungsevaluationen ihrer multilateralen Organisationen bis zu einem gewissen Grad systematisiert (vgl. OECD, 2013, S. 25).

Ab Mitte der 1980er war sich die DAC-Expert*innengruppe weitgehend darüber einig, dass bereits viel Wissen über „[...] what works and what does not“ existierte (OECD, 1988, S. 8).¹⁵³ Die Schwierigkeit, die bereits früh erkannt, aber bis dato vernachlässigt wurde, bestand eher darin, Evaluationsergebnisse in die Praxis zurückzuspielen (*Feedback*) und sie dort zu nutzen bzw. umzusetzen, demnach eine *Lern- bzw. Entwicklungsfunktion* zu etablieren (vgl. ebd.; Cailloux, 1988, S. 162).¹⁵⁴ Die Ansicht der Expert*innen-gruppe, dass Evaluation, inklusive dieser Funktionen, fester Bestandteil von EZ-Programmen sein sollte, wird in der Veröffentlichung „Methods and Procedures in Aid Evaluation – A Compendium of Donor Practice and Experience“ von 1986 deutlich, die weiträumig verbreitet und genutzt wurde. Neben Erläuterungen, wie Evaluationen im EZ-Kontext im Sinne „of both donors and recipient nations“ (OECD, 1986, S. 13) geplant und umgesetzt werden könnten, beinhaltete die Publikation auch *erste Definitionen von Evaluationsbegriffen* in Form eines *Glossars*, was Kjekhus (1991, S. 64) als eine wichtige Basis für ein gemeinsames Verständnis und für weitere Kooperationen bezeichnet. Sie war die erste

¹⁵² Z.B. die *Asiatische Entwicklungsbank* im Jahr 1981 und die *Welternährungsorganisation der Vereinten Nationen (FAO)* 1984 (Cracknell, 2002).

¹⁵³ 1984 wurde ein erster Evaluationsbericht mit dem Titel „Lessons of Experience“ vorgelegt.

¹⁵⁴ Die Expert*innengruppe veröffentlichte 1985 das Dokument „Good Feedback Practices“ (OECD, 1993). *Basil E. Cracknell* war damals Leiter der Evaluationsabteilung der englischen EZ-Agentur, Mitglied der Expert*innengruppe und federführend bei der Etablierung und Verbreitung des Feedback-Aspekts. S. hierzu ausführlich Cracknell (2002, S. 194ff.). In der Publikation „A Review of Donors’ Systems for Feedback from Aid Evaluation“ von 1990 wurde Feedback und die Nutzung von Ergebnissen zum zentralen Thema gemacht (vgl. OECD, 2013, S. 40).

in einer Reihe von Veröffentlichungen, die Grundsätze, Normen, Standards und Anleitungen hinsichtlich der Evaluierung von EZ bot (vgl. OECD, 2013, S. 32).

Cassens Studie „*Does Aid Work?*“ von 1986, die im Auftrag der Weltbank auf Programmebene durchgeführt wurde, wird als *wegweisend* bezeichnet und als erste Studie angesehen, „[...] to make judgement on aid effectiveness [...]“, indem wesentliche EZ-Aktivitäten überprüft und Ergebnisse der EZ-Evaluationen i.R. einer *Evaluationssynthese* zusammenführt wurden (Sasaki, 2006, S. 67). Die Ergebnisse der Studie sind komplex, führen den Autor allerdings zu einer positiven Beantwortung der Frage, während zeitgleich zahlreiche Bedenken und Mängel ausgesprochen werden (vgl. Cassen, 1986, S. 225). Die Studie steht zum einen für das zunehmende Interesse seitens der Wissenschaft an der Frage nach der Wirksamkeit von EZ (vgl. Cracknell, 2002, S. 45) und zum anderen für einen Wendepunkt in den 1980ern, da sie dazu führte, dass EZ-Organisationen sich zunehmend mit der Frage der Wirksamkeit ihrer Aktivitäten auseinandersetzten: „We are starting to adress issues of effectiveness in addition on those of efficiency“ (Cailloux, 1988, S. 162). Durch die Ergebnisse wurden die unterschiedlichen Zielsetzungen, Anforderungen und die damit verbundenen Herausforderungen hinsichtlich der Evaluationsaktivitäten in der EZ deutlich (Cracknell, 2002). Sie bestärkte die DAC-Expert*innen-Gruppe, den *allgemeinen Nutzen und die funktionelle Ausrichtung* von Evaluationsaktivitäten zu schärfen und voranzutreiben (vgl. ebd., S. 47).¹⁵⁵

Der Begriff des *Mikro-Makro-Paradoxons* wurde zu dieser Zeit zum ersten Mal publik und die EZ fiel in eine Glaubwürdigkeitskrise. Durch Mosleys Veröffentlichung (1986) wurde bekannt, dass Ergebnisse bei der Einzelevaluierung von Projekten (auf der Mikro-Ebene) meist positiv ausfielen, angestrebte Entwicklungserfolge auf der Makro-Ebene allerdings ausblieben (Neubert et al., 2014). Der Widerspruch löste grundsätzliche Debatten aus, die auch heute noch aktuell sind.¹⁵⁶

¹⁵⁵ Eine weitere Studie diskutiert das Für und Wider der damaligen ‚Entwicklungshilfe‘ (Riddell, 1987). Cracknell (2002, S. 47) bezeichnet Cassens und Riddels Beiträge als Meilensteine der Evaluation in der EZ auf ihrem Weg zur Reife zu einem eigenständigen Thema.

¹⁵⁶ S. auch Ausführungen unter Kapitel 4.4.

Um die Koordinierung von Studien zu verbessern (*Geberkoordinierung*) und den analytischen Umfang zu erweitern, wurden erste gemeinsame, *systemisch* ausgerichtete Evaluationsaktivitäten, sog. *Joint Evaluations* durchgeführt, bei denen mehrere Geber- und Partnerländer involviert waren. Diese Evaluationen zielten darauf ab, Evaluationskräfte hinsichtlich eines bestimmten Politikfelds, wie bspw. der Budgethilfe oder Friedensentwicklung, oder auch hinsichtlich eines Landes oder einer Region, zu bündeln und die Expertise hierzu zu erweitern (vgl. OECD, 2013, S. 17ff.). Bedingt durch die Hinwendung zu gemeinsamen Hilfsstrategien, sektorenbezogenen Programmen und anderen gemeinsamen Ansätzen in der EZ verlagerte sich das Interesse allgemein auf *Reviews* (bspw. länderbezogen) und *Evaluationssynthesen* zu spezifischen Themen, den sog. *cross-cutting issues* (OECD, 1993; vgl. OECD, 2013, S. 11ff.). Evaluationsberichte der Expert*innengruppe zu spezifischen Themen wie „[...] Sustainability, Impact on Women and Impact on Environment“ (ebd., S. 12) waren Ausgangspunkte für die OECD-Reihe „Selected Issues in Aid Evaluation“. *Nachhaltigkeit*, was durch die Einführung im *Brundtland-Report* der Vereinten Nationen 1987 zu einem wichtigen Gegenstand der Entwicklungsdebatte wurde, zog das Interesse der Expert*innengruppe auf sich. Der erste Evaluationsbericht zum Thema Nachhaltigkeit von 1988 war Grundlage für die Entwicklung der späteren DAC-Kriterien (vgl. ebd., S. 33). Zudem führten die amerikanische und kanadische EZ-Agentur Ende der 1980er erste *Meta-Evaluationen* durch, die zeigten, dass zahlreiche Evaluationen von nur unzureichender Qualität waren (vgl. Caspari, 2015, S. 153).

Um den Austausch und die Verbreitung von Evaluationsinformationen unter den OECD-Mitgliedern zu garantieren und den Zugang zu Evaluationsergebnissen zu vereinfachen (Kjekshus, 1991), wurde das *DAC Evaluation Report Inventory System* etabliert, in dem Berichte systematisiert, gespeichert und abgerufen werden konnten (OECD, 2013).¹⁵⁷

Auch einige *nichtstaatliche EZ-Organisationen* beschäftigten sich früh mit Evaluationsfragen, bspw. setzte das kirchliche Hilfswerk MISEREOR bereits in den 1960er Jahren *externe* Gutachtende zur Beratung und Evaluation ihrer Projekte ein (Nickolmann,

¹⁵⁷ Bis 1994 trugen 29 Agenturen ihre Bestände zusammen, die mehr als 1500 Evaluationsberichte und 945 geplante Evaluationen beinhalteten (OECD, 2013).

2006). Die Welthungerhilfe etablierte in den 1970er Jahren eine *interne* Stabsstelle für Evaluation (Guenther, 2006).¹⁵⁸

4.3 Phase III: Gezielte Umgestaltungen und komplexe Verhältnisse (1990er)

Mit Beginn der 1990er Jahre spricht Stockmann (2017, S. 36) von einem *zweiten Boom* der Evaluation, „[...] ausgelöst durch eine erhöhte staatliche Nachfrage in Nordamerika und Europa“ und „[...] durch eine Reihe von internationalen Entwicklungen [u.a. eine umfassende Verwaltungsmodernisierung] [...]“, der bis heute anhält (Stockmann und Meyer, 2017a, S. 69). Ab den 1990ern beginnt nach Wollmann (2017) die *Konsolidierungsphase*. Sie steht für eine *thematische, konzeptionelle und institutionelle Entfaltung und Entwicklung* der Evaluation (vgl. Wollmann, 2017, S. 43), aber auch für schnelle Wechsel hinsichtlich ihrer Rollenzuschreibung und Zielsetzung (*Kontrolle vs. Partizipation/gemeinsames Lernen*, vgl. Neubert, 2010, S. 142) sowie einer *Zunahme an Komplexität*:

„Der Paradigmen-Wechsel von einer Projektsicht hin zu Konditionalitäten und danach zu einer Perspektive, welche eher systemisch ausgerichtet war, wirkte sich auch auf die Evaluation aus. Systemische Betrachtungen und Bewertungen rückten ins Zentrum. Damit verbunden, erhöhte sich automatisch die Komplexität von Evaluationsvorhaben.“ (Stockmann & Borrmann, 2009, S. 34)

Zugleich vermehrte sich die *wissenschaftlich fundierte Kritik an der Wirksamkeit von EZ* sowie die kritische Auseinandersetzung mit den traditionellen EZ-Strukturen, was umfassende Anstrengungen auf nationaler und internationaler Ebene mit sich brachte, die bis in die 2000er hineinreichen (Faust & Neubert, 2010). Generell war die Kritik im Vergleich zu bisherigen zunehmend „konstruktiv, kenntnisreich und wissenschaftlich basiert formuliert“ und fand daher „oftmals auch Gehör in den Entwicklungsorganisationen“ (Faust & Messner, 2010, S. 271). Vor dem Hintergrund des wachsenden wirtschaftlichen, sozialen und politischen Drucks von außen, wie bspw. Haushaltsdefizite, ein

¹⁵⁸ Die Entwicklung von Evaluation bei Akteur*innen der nichtstaatlichen EZ in Deutschland stellt sich chronologisch anders dar als jene der staatlichen EZ. Bis zur Jahrtausendwende war die Evaluationssituation bei NGOs sehr divergent (Neubert et al., 2014). Auf eine ausführliche Darstellung und Aufnahme in die Chronologie wird daher verzichtet. S. hierzu weiterführend Neubert et al. (2014), Stockmann und Meyer (2017a) sowie die Systemprüfungen von Borrmann et al. (1999, 2001) und Borrmann und Stockmann (2009a, 2009b). Zu einer allgemeineren Betrachtung der EZ von NGOs und politischen Stiftungen s. Stockmann et al. (2016, S. 545ff).

wachsender Wettbewerb, mangelndes Vertrauen in die Regierung und der Ruf nach Legitimierung bzgl. der Ausgabe von Steuergeldern, unternahm zahlreiche OECD-Länder in den 1990er Jahren *umfassende Reformen*, auch i.R. ihrer Entwicklungspolitik (vgl. Binnendijk, 2000, S. 5; Faust & Neubert, 2010, S. 15).¹⁵⁹ Hier wurde der Fokus auf die Einbindung von Partner*innen, v.a. aber auf die Leistungsprobleme der EZ gelegt (Wirksamkeit, Effizienz, Qualität von Dienstleistungen): „[...] the development cooperation agencies have faced considerable external pressures to reform their management systems to become more effective and results-oriented“ (Binnendijk, 2000, S. 9). Der EZ-Mittelfluss war nunmehr „[...] an die Erfüllung meist expliziter, politischer und struktureller Bedingungen gebunden“, wodurch sich auch der Evaluationsgegenstand änderte: „Evaluiert wurde nun auch, ob Partnerregierungen den vereinbarten Konditionalitäten nachgekommen sind oder nicht“ (Bormann & Stockmann, 2009a, S. 34).¹⁶⁰

Aktivitäten vieler Geberländer legten im Folgenden ihren Schwerpunkt auf *größere Programme, Länderstrategien, sektorenübergreifende Vorhaben* und *Querschnittsthemen* (bspw. Armutsminderung, Menschenrechte, Gute Regierungsführung)¹⁶¹ statt wie bislang auf vereinzelte Projekte. Basierend auf der Erkenntnis, dass übergeordnete Systemzusammenhänge EZ-Erfolge offenbar drastisch beeinflussen, stellte sich kurz vor der Jahrtausendwende eine gewisse *Ernüchterung* in der internationalen EZ ein, die auch als *aid fatigue* bezeichnet wird (Binnendijk, 2000; Loewe, 2010; Nuscheler, 2008). Die seit Mitte der 1990er wachsende Zahl kritischer NGOs, die Arbeiten unabhängiger *Think Tanks* sowie seriöse Publikationen trugen dazu bei, die breitere Öffentlichkeit, v.a. die Fachöffentlichkeit in den Geber- und Nehmerländern, für grundlegende (Struktur-)Probleme der internationalen EZ zu sensibilisieren (Faust & Messner, 2010). Die *Defizite, traditioneller' EZ-Strukturen*, d.h. die organisatorische Aufstellung von EZ, stehen in einem

¹⁵⁹ Laut Faust und Messner (2010, S. 263) war das „grundsätzliche Informationsproblem des ‚EZ-Finanziers‘ über Verfahren und Effektivität der entwicklungspolitischen Zusammenarbeit [...] bis Mitte der 1990er Jahre besonders eklatant [...]“. Hinzu kamen auch Veränderungen auf internationaler Ebene, wie bspw. das Ende des Ost-West-Konflikts oder die Wirtschaftskrise in Asien, die gravierenden Einfluss auf die wirtschaftliche Entwicklung zahlreicher Länder hatten. Weitere Faktoren s. ebd. (S. 270f.)

¹⁶⁰ Der Druck blieb nicht ohne Konsequenzen: „Der Politikdialog wurde schwieriger und der internationale EZ-Mittelfluss brach ein“ (Stockmann & Bormann, 2009, S. 34).

¹⁶¹ Diese sektorenübergreifenden Ansätze wurden als sog. *SWAPs (Sector-Wide Approaches)* bezeichnet. Sie bilden gemeinsam mit den Ende der 1990er-Jahre entwickelten *Armutsbekämpfungsstrategien* eine wichtige Basis für *programm-basierte Ansätze (Programme-Based Approaches – PBA)*, die um die Jahrtausendwende zunehmend an Bedeutung gewannen (Klingebiel, Leiderer & Schmidt, 2007).

engen Zusammenhang mit dem Wirksamkeitsdefizit von EZ (Nuscheler, 2008; Faust & Messner, 2010). Zu solchen defizitären Grundmustern zählen „ein geringes Maß an *Ownership* auf der Nehmerseite¹⁶², eine stark zerklüftete Geberlandschaft¹⁶³“ sowie ein „hoher und ineffektiver Planungsapparat der EZ“.

Die *zunehmende Komplexität* der EZ-Strukturen sowie die Einführung *multidimensionaler Begriffe* bzw. *übergeordneter Ziele* (bspw. Armutsminderung) erschwerten die Evaluation (Cracknell, 2002; vgl. Terberger, 2011, S. 221) und zogen „[...] die Notwendigkeit der Entwicklung neuer Methoden“ nach sich (Neubert, 2010, S. 142). Diese ‚großen‘ Begriffe oder Ziele konnten nicht mehr mit einem einzigen numerischen Indikator wie z.B. Einkommenssteigerung ermittelt werden, sondern verlangten vielschichtig angelegte Untersuchungskriterien, die nur qualitativ beschreibbar schienen. „Seither wurden in der EZ nach Methoden zur Erfassung komplexer Zielkonzepte gesucht“ (ebd., S. 142).

Als der wichtigste Meilenstein für die EZ-Evaluation in den 1990er Jahren können die „*DAC Principles on Evaluation of Development Assistance*“ (OECD, 1991) betitelt werden (vgl. OECD, 2013, S. 40). Die Grundprinzipien der *Unabhängigkeit, Unparteilichkeit, Glaubwürdigkeit, Transparenz, Nützlichkeit, Aktualität und Inklusivität* waren nicht neu, aber es war wichtig, sie als Prinzipien zu etablieren. Auch die *Definition von Evaluation* war ein wichtiger Baustein. Fünf darin beinhaltete Wörter hatten einen tiefgreifenden Einfluss auf die Bewertung von EZ: *Relevanz, Effektivität, Effizienz, Wirkungen und Nachhaltigkeit* (ebd.). Diese Evaluationskriterien werden seither aktiv genutzt und sind in vielen EZ-Organisationen Teil des Projektmanagements und damit auch der Berichterstattung.

¹⁶² Die stärkere Beteiligung von Zielgruppen in Projektabläufe wird seit Ende der 1980er von einigen EZ-Akteur*innen forciert, das *Ownership*-Prinzip seit den 1990ern seitens zahlreicher EZ-Organisationen immer wieder propagiert, allerdings kaum umgesetzt (Faust & Messner, 2010).

¹⁶³ Neben bilateralen Gebenden, globalen und regionalen Finanzinstitutionen, UN-Agenturen sowie einer steigenden Anzahl globaler Fonds kommen sog. „neue Geber“ wie China, Indien, Südkorea oder die Türkei hinzu, die keine OECD-Länder sind. Als ein Länderbeispiel für die Überkomplexität wird Uganda genannt, das (Stand 2005/2006) von ca. 40 Geberorganisationen mit Entwicklungshilfe versorgt wird und die Regierung nach eigenen Angaben mit 684 unterschiedlichen Instrumenten der EZ umgehen muss (Faust & Messner, 2010, S. 264). Fues (2010) gibt einen Überblick über das komplexe UN Entwicklungssystem, beschreibt Stärken und Schwächen hinsichtlich der Wirksamkeit der UN-EZ und geht auf Reformbemühungen ein.

Nachhaltigkeit als Evaluationskriterium rückte trotz vorangegangener Debatten „[...] erst ab Ende der 1980er-Jahre in den Fokus der deutschen Wirksamkeitsdebatte“ (Stockmann & Gaebe, 1993 zitiert in Noltze, Euler & Verspohl, 2018a, S. 9). Eine Empfehlung der DAC-Expert*innengruppe gab 1986 den Anstoß, Nachhaltigkeit bei der Erfolgsbewertung von Projekten zu berücksichtigen, was zunächst durch *ex-post Evaluationen* bei der KfW, später auch vermehrt durch die GTZ umgesetzt wurde (Stockmann & Gaebe, 1993; Noltze et al., 2018a). Doch erst Ende der 1990er wurden GTZ und KfW seitens des BMZ dazu angehalten, die Kriterien i.R. ihrer Projektevaluationen anzuwenden (vgl. Zintl, 2006, S. 42). Der Lehrstuhl für Soziologie unter der Leitung *Reinhard Stockmanns* an der Universität des Saarlandes widmete sich dem Nachhaltigkeitsthema intensiv und bereits früh in den 1990ern. Neben Studien zur Nachhaltigkeit von Berufsbildungsprojekten in der EZ (Stockmann, 1992, 1996) und zur Langfristigkeit von EZ-Projekten allgemein (Stockmann & Gaebe, 1993) wurde in Kooperation mit dem BMZ eine Studie durchgeführt, bei der 32 Vorhaben mittels *ex-post-Evaluationen* durch eine große Zahl von Gutachtenden u.a. hinsichtlich ihrer langfristigen Wirksamkeit untersucht wurden (vgl. BMZ, 2000). Die Studie stellte „den Auslöser für die systematische Auseinandersetzung mit Nachhaltigkeit als Erfolgskriterium der deutschen EZ [dar]“ (Noltze et al., 2018a, S. 9). Caspari (2004) nutzte später die Ergebnisse der BMZ-Studie (2000) für eine weitere *Evaluationssynthese* zur Evaluation der Nachhaltigkeit von EZ, betrachtete die Evaluations- und Bewertungspraxis und aggregierte die Ergebnisse der Studien inhaltlich (Noltze et al., 2018a). „Die damit einhergehende Debatte machte deutlich, dass Nachhaltigkeit zu diesem Zeitpunkt in der deutschen EZ sehr heterogen verstanden und bewertet wurde, was die Möglichkeiten einer inhaltlichen Querschnittsauswertung deutlich einschränkte“ (ebd., S. 9). Die Studien haben dazu beigetragen, den Nachhaltigkeitsbegriff zu schärfen und das Konzept der Nachhaltigkeit analytisch zu erfassen.

Als *Schlüsselement* der damaligen Reformen zur Politik- und Verwaltungsmodernisierung, wie dem *New Public Management*,¹⁶⁴ wird der *theoriegeleitete Ansatz des Results*

¹⁶⁴ *NPM* ist ein Konzept zur Verwaltungsmodernisierung, das seit Ende der 1980er international und seit Anfang der 1990er in Deutschland Einzug fand (Wollmann, 2017). Es „[...] verfolgt das Ziel, die Leistungen und den Prozess der Leistungserbringung in der öffentlichen Verwaltung zu verbessern“ (Stockmann, 2006, S. 61).

*Based Management (RBM)*¹⁶⁵ betitelt, der als Ansatz des *Performance Managements (PM)* sowohl eine strategische Planung und eine Erfolgsbewertung beinhaltet und starke Ähnlichkeiten mit dem *Logical Framework-Ansatz* hat. Die Leistungssteigerung („bessere Ergebnisse zu erzielen“) stellen die zentrale Ausrichtung dar (Binnendijk, 2000).¹⁶⁶ Da Evaluation in einigen EZ-Agenturen bereits fester Bestandteil war bevor PM-Ansätze eingeführt wurden, die nicht zwangsweise Evaluation beinhalteten, war das Verhältnis zueinander nicht immer eindeutig.¹⁶⁷ Einerseits wurde Evaluation eine komplett gesonderte Funktion und unabhängige Rolle zugesprochen, andererseits als ein integraler Bestandteil von PM gesehen. Der Mittelweg sieht Evaluation in das PM integriert, allerdings mit einer eigenständigen und speziellen Funktion:

„Evaluation is viewed as complementary to – and in some respects superior to – other routine performance measurement techniques. For example, evaluation allows for more in-depth study of program performance, can analyze causes and effects in detail, can offer recommendations, or may assess performance issues normally too difficult, expensive or long-term to assess through on-going monitoring“ (Binnendijk, 2000, S. 8).

Dieser ‚Mittelweg‘ wurde in der Praxis zunehmend von EZ-Agenturen akzeptiert und übernommen. In der „Best Practice Guideline for Evaluation“ der OECD von 1998 heißt es: „[...] evaluation must be part of a wider performance management framework“ (OECD, 1998).

Basierend auf vorangegangenen Debatten fand ein neues Verständnis von Partnerschaften allmählich Einzug in die EZ-Arbeit und *partizipative Evaluationsansätze* trafen „[...] im Laufe der 1990er Jahre auf breitere Zustimmung in der EZ-Praxis“ (Neubert, 2010, S. 141). Erst jetzt öffneten sich viele EZ-Organisationen für die Idee, Partner*innen neben der Planung und Durchführung von Vorhaben auch in die Evaluation einzubeziehen.

¹⁶⁵ Dem Ansatz wurden viele Namen gegeben (u.a. *strategic management, performance-based management*) (Swiss, 2005). Basierend auf genutzter Literatur wird der Begriff *Results Based Management (RBM)* verwendet.

¹⁶⁶ Die Veröffentlichung von Binnendijk (2000) wurde aufgrund der raschen Etablierung und zunehmend wichtigen Rolle von RBM in den Entwicklungsagenturen seitens der OECD in Auftrag gegeben (OECD, 2013).

¹⁶⁷ Die neuen PM-Techniken wurden u.a. entwickelt, um Schwächen von Evaluationen, bspw., dass Evaluationsergebnisse verhältnismäßig wenig aussagekräftig zu den Ausgaben standen, auszugleichen. Evaluation wurde häufig als Aufgabe externer Expert*innen betrachtet, während PM eine wesentliche Aufgabe der Organisation selbst war. Das ungeklärte Verhältnis barg Doppelarbeit, Verwirrungen und Spannungen zwischen internen Organisationsabteilungen und professionellen Gruppierungen (vgl. Binnendijk, 2000, S. 7f.).

Hierbei wurden konkrete und einfach durchführbare Ansätze wie das *Participatory Impact Monitoring (PIM)* entwickelt, die sich an Instrumenten des PRA bedienen (Bliss, Neubert & Heinz, 2007).¹⁶⁸ PIM wurde zu Beginn der 1990er im Auftrag der GTZ entwickelt und sollte eine Alternative zur ZOPP darstellen (Gohl, 2010). *MAPP (Method for Impact Assessment of Programs and Projects)* ist ein weiterer partizipativer Ansatz der Wirkungsanalyse, der Ende der 1990er am *Deutschen Institut für Entwicklungspolitik (DIE)* entwickelt und von deutschen EZ-Agenturen v.a. in Afrika und Lateinamerika eingesetzt wurde (Neubert & Müller, 2010; Neubert, 2010). Der Ansatz ist kontextorientiert und im Kern qualitativ, allerdings enthält er einen Quantifizierungsschritt, durch den Ergebnisse aggregiert werden können (Neubert, 2010).¹⁶⁹ Hinsichtlich der *Stärkung der Evaluationskapazitäten in den Partnerländern* fand 1990 ein erstes regionales OECD-Evaluationsseminar in Abidjan/Elfenbeinküste statt. Der Erfolg des Seminars führte zu zahlreichen regionalen Folgeveranstaltungen. Das Format wurde jedoch Ende der 1990er eingestellt (OECD, 2013).¹⁷⁰

Zur neuen Bewegung der *sektorenübergreifenden Programme* wurden erste richtungsweisende Beiträge veröffentlicht (u.a. Brown & Foster, 2001; Foster, Brown & Conway, 2000; Riddell, 2002), die auch Überlegungen zur Evaluationspraxis beinhalteten. Einige EZ-Organisationen experimentierten bei *Programmevaluationen auf Länderebene* mit verschiedenen Ansätzen und führten zahlreiche Studien durch (OECD, 2013).¹⁷¹ Dennoch wurden nach wie vor noch zahlreiche *Evaluationen von Einzelvorhaben* durchge-

¹⁶⁸ Ausgehend vom PRA zeigen Bliss et al. (2007) verwandte und abgeleitete Ansätze auf (bspw. *Participatory Poverty Assessment, Participatory Learning Methods* und *ZOPP*) und geben einen umfangreichen Überblick über partizipative Methoden und Ansätze sowie konkrete Praxistipps für die Durchführung.

¹⁶⁹ Zur detaillierten Aus- und Durchführung des *PIM-Ansatzes* s. Gohl (2010, S. 120ff.), zu *MAPP* s. Neubert (2010, S. 153ff.), Neubert und Müller (2010, S. 88ff.)

¹⁷⁰ Im Anschluss wurde seitens der DAC-Mitglieder, multilateralen Entwicklungsbanken und dem Entwicklungsprogramm der Vereinten Nationen (UNDP) der entstandene Dialog fortgeführt, allerdings über verschiedene Formate. Z.B. wurden gemeinsame Evaluationen durchgeführt, finanziert oder kleinere regionale Seminare und Konferenzen abgehalten (vgl. OECD, 2013, S. 29).

¹⁷¹ Einen Überblick zum Status Quo der damaligen Programmevaluation bietet das Handbuch der Weltbank (Valadez & Bamberger, 1994). 1994 und 1999 fanden zwei OECD-Workshops statt, um hierzu Erfahrungen und Stärken und Schwächen verschiedener Ansätze identifizieren zu können. Resultierend daraus wurde die Publikation „Evaluating Country Programmes“ (OECD, 1999) veröffentlicht (OECD, 2013).

führt und deren Ergebnisse im Anschluss mittels *Evaluationssynthesen* zusammengeführt, bspw. wie die Weltbank in sog. *Länderportfolios* (Bliss, 2006).¹⁷² In Deutschland wurden Evaluationen zu Länderprogrammen erst in den 2000ern erprobt und durchgeführt. Dafür gab es aber gelegentlich größere *Evaluationssynthesen*, wie ab 1994 die jährlichen der GTZ zu allen laufenden und abgeschlossenen Projekten (vgl. Reuber, 2006, S. 14f., 80).¹⁷³ An groß angelegten *Reviews* wie bspw. zum Genozid in Ruanda beteiligten sich mehrere OECD-Länder, UN-Organisationen und NGOs.¹⁷⁴ Die *Joint Evaluations* wurden zunehmend populärer (OECD, 2013). Mit der Veröffentlichung „Donor Experience with Joint Evaluations – A Typology and Lessons Learned“ von 1998, die zeigte, dass gemeinsame Evaluationen die Entwicklung einer gemeinsamen Evaluationskultur unterstützen und damit Partnerschaften zwischen gebenden und Empfangenden stärken und das Bewusstsein für verschiedene Ansätze und Methoden schärfen konnten, stieg die Anzahl an *Joint Evaluations* weiter an.¹⁷⁵ Zudem wurden „[...] in den 1990er Jahren international mehrere durchaus umfangreiche [*Meta-*]Evaluationsen umgesetzt [...]“, die wiederum zeigten, dass ein Großteil der Evaluationen qualitativ umfangreiche Mängel aufwiesen (Caspari, 2015, S. 153f.). Der „[...] Prozess ebte [allerdings] mit Beginn der Diskussion um evidenzbasierte Evaluationen [...] wieder ab“ (ebd., S. 154). Das Entwicklungsprogramm der Vereinten Nationen (*United Nations Development Programme – UNDP*) veröffentlichte ab dato die *Human Development Reports*, an deren Analysen zunehmend auch lokale Wissenschaftler*innen beteiligt waren (Fues, 2010).¹⁷⁶ Die Berichte enthalten den *Index der menschlichen Entwicklung* (engl.: *Human Development*

¹⁷² D.h. alle Projekte, die durch die Bank in einem Jahr oder in einem größeren Zeitraum finanziert wurden (Bliss, 2006). Auch die DAC-Expert*innengruppe fokussierte sich auf Synthesen, bspw. eine Studie zum Thema Umwelt, die 1997 angelegt, 70 Evaluationen zum Thema vereint und im Jahr 2000 publiziert wurde (OECD, 2013).

¹⁷³ Bis 1998 waren in Deutschland die Jahresauswertungen die umfangreichsten Evaluationen. 1996 und 1997 wurden hierbei rund 110 Einzelevaluationen berücksichtigt (vgl. Bliss, 2006, S. 15). Zur damaligen Evaluationspraxis der GTZ im Detail, s. Reuber (2006).

¹⁷⁴ Laut OECD (2013, S. 18f.) hatte diese Evaluation erhebliche Auswirkungen auf den humanitären Sektor, da sie u.a. eine Reihe von Initiativen zur Qualitätssicherung auslöste, wie bspw. das *Active Learning Network on Accountability and Performance*. Wenn einem Thema wie in diesem Falle eine ganz besondere Wichtigkeit zugeschrieben wurde, wurden die Studien auch *DAC Evaluationen* genannt (ebd.).

¹⁷⁵ S. aufgelistete Beispiele OECD (2013, S. 17).

¹⁷⁶ Die Berichte sind öffentlich zugänglich unter: <http://www.hdr.undp.org/en/global-reports>.

Index, HDI), der die durchschnittlichen Landeswerte in grundlegende Bereiche der menschlichen Entwicklung erfasst.¹⁷⁷

Die 1997 gegründete *Deutsche Gesellschaft für Evaluation (DeGEval)* zählt zu den ersten nationalen Fachverbänden in Europa. „Vor 1995 existierten lediglich fünf regionale und/oder nationale Evaluationsorganisationen [...]“ (Rodríguez-Bilella & Lucero, 2017, S. 55).¹⁷⁸ Laut Wollmann (2017, S. 45) war dies ein wichtiger Schritt der „institutionellen und professionellen Konsolidierung“ der Evaluationsforschung in Deutschland. Inzwischen haben sich innerhalb des Verbandes 16 Arbeitskreise und ein Nachwuchsnetzwerk entwickelt, darunter auch der *Arbeitskreis Entwicklungspolitik und Humanitäre Hilfe* (DeGEval, 2018, 6. August).¹⁷⁹ Der Arbeitskreis ist einer der ältesten des Netzwerks, da „kaum ein anderes Politikressort [...] so oft evaluiert [wird und wurde] wie das der Entwicklungspolitik – und gleichzeitig so umstritten [ist]“ (Mayring et al., 2017, S. 17). Als ein weiterer wichtiger Schritt bzgl. der Institutionalisierung von Evaluation in der EZ soll an dieser Stelle noch die Gründung der *African Evaluation Association (AfREA)* im Jahr 1999 erwähnt werden. Sie wurde seitens zahlreicher Mitglieder der DAC-Expert*innen-gruppe und UN-Organisationen unterstützt und ist ein wichtiger Meilenstein hinsichtlich des Auf- und Ausbaus von Evaluationskapazitäten in Afrika (OECD, 2013).¹⁸⁰

¹⁷⁷ Wie bspw. die Lebenserwartung bei der Geburt, das Bildungsniveau sowie das Pro-Kopf-Einkommen. Aus einer Vielzahl solcher Einzelindikatoren wird eine Rangliste errechnet. Mittels der Liste kann der durchschnittliche Entwicklungsstand eines Landes abgeleitet werden (BMZ, 2010-2019).

¹⁷⁸ Die *American Evaluation Association (AEA)*, Gründung 1986), *Australasian Evaluation Society (AES)*, 1987), *Canadian Evaluation Society (CES)*, 1979), *Central American Evaluation Association (ACE)*, Gründung unbekannt) und *European Evaluation Society (EES)*, 1992) (Rodríguez-Bilella & Lucero, 2017, S. 56ff.).

¹⁷⁹ 2006 wurde die Vereinigung in *DeGEval – Gesellschaft für Evaluation* umbenannt, um auch österreichische Mitglieder mit einzubeziehen (Mayring et al., 2017). Eine Auflistung der Hauptziele, Publikationen und eine Beschreibung der Arbeitskreise bieten Böttcher, Kerlen, Maats, Schwab und Sheikh (2014, S. 208ff.). Mayring et al. (2017) beschreiben die Geschichte und die Entwicklung der Themenbereiche und Arbeitskreise.

¹⁸⁰ Bei seiner Gründung vereinte der Verband sechs nationale Fachverbände, heute sind es 35. Bedeutende Schritte waren die Entwicklung der *African Evaluation Guidelines (AfREA)*, 2006/07) und des „Africa-rooted“-Evaluationsansatzes (s. Chilisa, 31. August 2015) (vgl. Rodríguez-Bilella & Lucero, 2017, S. 57).

4.4 Phase IV: Konsolidierung und Institutionalisierung (2000-2015)¹⁸¹

Zur Jahrtausendwende vollzog sich ein grundlegender *Wandel hinsichtlich der Wirksamkeitsdebatte* in der EZ (vgl. Ashoff, 2010, S. 47; Klingebiel, 2013, S. 58ff., 2017, S. 453; Neubert et al., 2014, S. 63; Oberndörfer et al., 2010, S. 9; Roxin, 2016, S. 12). Mit Beginn und im Laufe der Nullerjahre wurde eine Reihe *internationaler Konsense* beschlossen und Initiativen gestartet, die die EZ insgesamt sowie die Evaluation von EZ nachhaltig beeinflussten (Borrmann & Stockmann, 2009a). Nachdem die zunehmende Skepsis der Öffentlichkeit lange ignoriert „[...] oder ihr mit bloßen Argumenten statt Evidenz begegnet wurde [...]“, wurde der Druck, Erfolge nachzuweisen, auch auf politischer und fachlicher Ebene erhöht (Roxin, 2016, S. 12). Im Zuge dessen kam es zu einem regelrechten „Boom in der internationalen Wirkungsforschung, [allerdings] mit kontroversen Erkenntnissen“ (Nuscheler, 2008, S. 23).¹⁸² Da die Ausdifferenzierung und Professionalisierung sowie die Frequenz von Evaluationsaktivitäten zunahmen, werden ab hier aus Gründen der besseren Übersicht Zwischenüberschriften genutzt.

Kontext und Meilensteine

Die *Millenniums-Deklaration* und die *Millenniums-Entwicklungsziele (Millennium Development Goals - MDGs)* der Vereinten Nationen waren ein Signal des Aufbruchs und wichtige Meilensteine in der Reform-Dynamik der EZ. Als anerkannter und verbindlicher Referenzrahmen legten die MDGs zum ersten Mal allgemeingültige Ziele fest, die überprüfbar und mit klar formulierten, quantitativen sowie zeitlich begrenzten Erfolgsindikatoren hinterlegt waren. Die internationale Gebergemeinschaft einigte sich damit auf einen „relativ konkreten Zielrahmen“ mit einem entsprechenden, für die EZ-Akteur*innen verbindlichen Monitoring-System (Borrmann & Stockmann, 2009a, S. 35; Brand, 2006; Oberndörfer et al., 2010).¹⁸³ Nachdem i.R. des *Monterrey-Konsens* 2002 eine massive Erhöhung der EZ-Mittel beschlossen wurde¹⁸⁴ und die Geberländer sich in der

¹⁸¹ Aufgrund der Laufzeit der MDGs und damit verbundene Evaluationsaufgaben, reicht dieses Teilkapitel bis 2015.

¹⁸² Ashoff (2010, S. 29f.) gibt einen umfassenden Überblick der „Bandbreite der Kritik“ hinsichtlich der Wirksamkeit der Entwicklungspolitik im Allgemeinen.

¹⁸³ Loewe (2010) beschreibt die Entstehungsgeschichte der MDGs und erörtert ihre Wirkungsmechanismen, Risiken, Herausforderungen und Konsequenzen für die EZ.

¹⁸⁴ Auch als ‚*Big Push*‘ bezeichnet, s. hierzu Caspari (2009, S. 185) und Nuscheler (2008, S. 20f).

Marrakesch-Deklaration u.a. auf eine „Steuerung nach Resultaten“ einigten, wurde 2005 die *Paris Declaration on Aid Effectiveness* verabschiedet (Borrmann & Stockmann, 2009a, S. 35). Die Kernprinzipien der Agenda greifen zentrale Kritikpunkte am „alten EZ-Regime“ auf und beschäftigt sich damit, „wie“ EZ künftig „wirksamer“ durchgeführt werden sollte (vgl. Ashoff, 2010, S. 49; Nuscheler, 2008, S. 17; Roxin, 2016, S. 12). Die Vereinbarungen von Paris wurden in weiteren Foren zur *Aid Effectiveness* weiterentwickelt und 2011 von 100 Nationen unterzeichnet.¹⁸⁵ Die Reihe an Beschlüssen, die seit 2000 gefasst wurden, bildet zusammen die *internationale Aid Effectiveness Agenda*. Die Aspekte und Prinzipien der Agenda sind nicht völlig neu, „[...] sondern haben sich sukzessive aus den Erfahrungen früherer Jahre entwickelt“ (Kranz-Plote, 2006, S. 96; vgl. auch Ashoff, 2010, S. 50). Dennoch hatten die Beschlüsse große Auswirkungen auf die Umsetzungsmodalitäten der EZ¹⁸⁶ und beinhalten eine weitere *Zunahme an Komplexität und methodischen Herausforderungen*, was die Evaluation anbelangt (Stockmann & Borrmann, 2009, S. 266), bspw. höhere Anforderungen an Wirkungsmessungen (und Ursachenanalysen) durch die Verschiebung der Entwicklungsziele auf höhere Aggregatenebenen¹⁸⁷ und komplexere Evaluationsverfahren, inklusive anspruchsvolleren Designs und Methoden. Des Weiteren wurde durch die zunehmende Vernetzung der Geberorganisationen, um gemeinsam aufgelegte Programme evaluieren zu können, die *Verein-*

¹⁸⁵ Es fanden vier *High Level Fora on Aid Effectiveness* statt: Rom (2003), Paris (2005), Accra (2008), Busan (2011) (OECD, 2019a). Details zu den Foren und zur Evaluation der Paris Deklaration s. Dabelstein und Patton (2015).

¹⁸⁶ Bspw. Zunahme sektoraler Budgethilfen, weniger Projektmittel, komplexe bi- und multilaterale Ko-Finanzierungen, Programmhilfen, Korbfinanzierungen mit den wichtigsten Akteur*innen eines Sektors etc. (Geberkooperation und -harmonisierung). Zu nationalen Umsetzungsstrategien hinsichtlich der MDGs und angepassten Leitlinien zur TZ und FZ s. Stockmann und Borrmann (2006, S. 266ff). Zum Reformbedarf der deutschen EZ-Instrumente s. auch Scholz und Wolff (2010).

¹⁸⁷ Eine ‚einfache‘ Ergänzung der bestehenden M&E-Systeme auf Projekt- und Programmebene um eine aggregierte und MDG-orientierte Gesamtperspektive gestaltete sich als schwierig (Brand, 2006).

heitlichung der Evaluationsterminologie, -standards, -konzepte und -verfahren zwingend notwendig. Darüber hinaus setzte die Übertragung der Verantwortlichkeit für Evaluation in die Partnerländer die *Entwicklung von Evaluationskapazitäten* voraus.¹⁸⁸

Generell wurde in der EZ bis dato zwar viel evaluiert, doch existieren zahlreiche Schwachpunkte, die sich nach Faust (2010) in organisatorisch-institutionelle und methodische unterteilen lassen: Die *organisatorisch-institutionelle Kritik* betont die *Vernachlässigung der Wirkungen* i.R. der Evaluationspraxis (Fokus auf Inputs und Output), dem *isolierten Vorgehen der Geberorganisationen* sowie die *begrenzte Unabhängigkeit*, „[...] was die Ergebnisse zu Gunsten positiver Bewertungen verzerrt“ (ebd., S. 43). White und Bamberger (2008, S. 1) sprechen von einem allgemeinen *positiven Bias* hinsichtlich der Projektevaluation, der auf die Zensur seitens der Organisation oder beauftragter Evaluierender zurückzuführen ist. Die *methodische Kritik* bezieht sich auf die Art und Weise, wie Wirkungsmessungen durchgeführt wurden: mangelhafte Datenerhebungen, fehlende Baseline-Studien, Erklärungsmangel hinsichtlich des Kontrafaktischem und ein sog. Selektionsbias bei der Projektauswahl, d.h. nur besonders erfolgreiche Projekte wurden evaluiert (vgl. Faust, 2010, S. 43; Roxin, 2016, S. 13f.). Die Reformdynamik in der internationalen EZ mündete in den Nullerjahren sowohl in *inhaltlich-methodischen als auch in institutionell-organisatorischen Bemühungen*, die im Folgenden weiter erläutert werden sollen.

Standardisierung

Nach langer Vorarbeit wurde 2002 das „Glossary of Key Terms in Evaluation and Results Based Management“ (OECD, 2009) mit Definitionen von und rund um Evaluation veröffentlicht. Die Einigung auf ein Glossar war schwierig, bspw. existierten bzgl. des Begriffs *Impact Evaluation* über 20 verschiedene Definitionen. Es existiert mittlerweile in 16 ver-

¹⁸⁸ Je mehr bspw. die Prinzipien von Partnerorientierung und Geberkooperation umgesetzt werden, desto weniger ist es sinnvoll und möglich, positive Wirkungen bzw. ‚Erfolge‘ einzelnen EZ-Organisationen zuzuschreiben (Kranz-Plote, 2006). In der internationalen Debatte besteht zu diesem Zeitpunkt Einigkeit, „[...] dass eine exakte Quantifizierung und Zuordnung des Wirkungsbeitrags einzelner Geber [...], weder sinnvoll noch möglich ist“ sowie „[...] dass die Ebene des einzelnen Vorhabens bei einer Wirkungsbeurteilung, die auch Bezug auf solche übergeordneten Entwicklungsziele nehmen soll, nicht ausreicht“ (ebd., S. 92). Weitere Beiträge in Nuscheler und Roth (2006) beleuchten die MDGs kritisch aus verschiedenen Perspektiven.

schiedenen Sprachen und ist womöglich das weit verbreitetste Produkt der DAC-Expert*innengruppe (OECD, 2013). Die DAC-Qualitätsstandards „for Development Evaluation“ (OECD, 2010b) unterlagen einem ähnlich langwierigen Prozess vom Konsens ihrer Notwendigkeit bis zur Veröffentlichung. Im Juni 2010 wurden alle Schlüsseldokumente der Expert*innengruppe in einer Sammlung mit dem Titel „Summary of Key Norms and Standards“ (OECD, 2010a) veröffentlicht (OECD, 2013). Die *Definition von Evaluation* umfasst nach Dahler-Larsen (2017) *alle typischen Definitionsbestandteile* und kommt den allgemein anerkannten und akzeptierten Definitionen in der Evaluationsforschung recht nahe (Caspari, 2004):

„[Evaluation/Evaluierung ist eine] systematische und objektive Beurteilung eines laufenden oder abgeschlossenen Projekts oder Programms bzw. einer Politikmaßnahme und ihrer Konzeption, Umsetzung und Ergebnisse. Ziel ist es, die Relevanz und Verwirklichung von Zielen zu ermitteln sowie ihre Effizienz für die Entwicklung, ihre Wirksamkeit, längerfristige Wirkung und Nachhaltigkeit zu bestimmen. Eine Evaluierung sollte glaubwürdige und nützliche Informationen liefern, die die Einbeziehung gewonnener Erkenntnisse in den Entscheidungsfindungsprozess von Empfängern und Gebern ermöglichen. Ferner bezeichnet der Begriff „Evaluierung“ einen Prozess, in dem der Wert oder die Signifikanz einer Aktivität, einer Politik oder eines Programms bestimmt werden, mit anderen Worten eine möglichst systematische und objektive Bewertung einer geplanten, laufenden oder bereits abgeschlossenen Entwicklungsmaßnahme.“ (OECD, 2009, S. 27f.)

Monitoring definiert die OECD (2009, S. 35f.) als einen „kontinuierliche(n) Prozess der systematischen Datensammlung zu spezifischen Indikatoren, um Leitung und Hauptbeteiligte eines laufenden EZ-Projekts über die erzielten Fortschritte und die erreichten Ziele sowie über die Verwendung bereitgestellter Mittel zu informieren“. Eine *Review* zur Evaluationspraxis der DAC-Mitglieder (OECD, 2010c) zeigte, dass Standards wie die Evaluations- und Qualitätskriterien weitgreifend genutzt werden und sich Evaluationsrichtlinien fast aller Mitglieder auf die DAC-Prinzipien und Evaluationsstandards beziehen. Allerdings weist Caspari (2004, S. 15f.) immer noch auf bestehende *Definitionsschwierigkeiten* hin, die Borrmann et al. (1999, S. 13) auch als „babylonisches Sprachgewirr“ bezeichnen: Häufig wird hier Evaluation mit Begriffen wie *entwicklungspolitische Erfolgskontrolle*, *Wirkungskontrolle*, *Wirkungsuntersuchung*, *Wirksamkeitsuntersuchung*, *Inspektion*, *Fortschrittskontrolle*, *Effektivitäts- und Effizienzanalyse* gleichgesetzt (Caspari, 2004).

Die DeGEval veröffentlichte 2001 in Anlehnung an die US-amerikanischen Standards und die DAC-Prinzipien eigene *Standards für Evaluation zur Qualitätssicherung* (Bliss, 2006).¹⁸⁹ Anders als bei den deutschen Evaluationsstandards (vgl. Kap. 3.1.4) wird die *Unabhängigkeit* von Evaluierenden in den *DAC Principles*¹⁹⁰ mehrfach und in den *DAC Evaluation Quality Standards* explizit als *Independence of evaluators vis-à-vis stakeholders* erwähnt:

„Evaluators are independent from the development intervention, including its policy, operations and management functions, as well as intended beneficiaries. Possible conflicts of interest are addressed openly and honestly. The evaluation team is able to work freely and without interference. It is assured of co-operation and access to all relevant information“ (OECD, 2010a, S. 25).

Erstes wichtiges Ergebnis der deutschen Arbeitsgruppe *Evaluierung aus einem Guss*¹⁹¹ war die Entwicklung von „Evaluierungskriterien für die deutsche bilaterale Entwicklungszusammenarbeit“ in Anlehnung an die DAC-Kriterien mit dem Ziel, ein gemeinsames Verständnis von Evaluation herzustellen. Seit 2006 werden alle Evaluationen, die vom BMZ beauftragt werden, anhand dieser *Orientierungslinien* bewertet, auch jene, die von nicht-staatlichen Organisationen durchgeführt werden (Borrmann & Stockmann, 2009a; Neubert et al., 2014; Stockmann & Borrmann, 2009; Zintl, 2006, 2009;).¹⁹²

Da *Menschenrechte* und die *Gleichstellung von Geschlechtern* bislang i.R. von Evaluationen zu wenig Beachtung fanden, gründete die *United Nations Evaluation Group (UNEG)* eine Arbeitsgruppe zur systematischen Einbindung der Themen und publizierte ein umfangreiches Handbuch (vgl. UNEG, 2011).

¹⁸⁹ Die Standards wurden 2016 in einer überarbeiteten Fassung neu veröffentlicht (DeGEval, 2016).

¹⁹⁰ „Impartiality and Independence: The evaluation process should be impartial and independent from the process concerned with policy making, and the delivery and management of development assistance. (Para. 11); Impartiality contributes to the credibility of evaluation and the avoidance of bias in findings, analyses and conclusions. Independence provides legitimacy to evaluation and reduces the potential for conflict of interest which could arise if policy makers and managers were solely responsible for evaluating their own activities. (Para. 12); Impartiality and independence will best be achieved by separating the evaluation function from the line management responsible for planning and managing development assistance. [...]“ (OECD, 2010a, S. 8)

¹⁹¹ S. Ausführungen unter der Überschrift *Institutionalisierung und Professionalisierung*.

¹⁹² Die „Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen“ beinhaltet Evaluationskriterien nach DAC (Relevanz, Effektivität, Effizienz, Entwicklungspolitische Wirkungen und Nachhaltigkeit) sowie zusätzliche Kriterien des BMZ (Kohärenz, Komplementarität & Koordination) und jeweils dazugehörige Prüffragen. Zudem wurde ein einheitlicher Sprachgebrauch für die DAC-Wirkungskette festgelegt (vgl. BMZ, 2006).

Wegweisende Studien

Die Studie „*When will we ever learn? Improving lives through Impact Evaluation*“ der *Evaluation Gap Working Group* des *Center for Global Development (CGD)*¹⁹³ (Savedoff, Levine & Birdsall, 2006) sorgte für viel Aufruhr und stieß zahlreiche Veränderungsprozesse an (vgl. bspw. Borrmann & Stockmann, 2009a, S. 49f; Caspari, 2008, S. 137; 2009, S. 188; Caspari & Barbu, 2008, S. 2; Faust, 2010, S. 41; OECD, 2013, S. 37; Roxin, 2016, S. 14; White & Bamberger, 2008, S. 3). Hier wird einleitend festgestellt:

„Yet after decades in which development agencies have disbursed billions of dollars for social programs, and developing country governments and nongovernmental organizations (NGOs) have spent hundreds of billions more, it is deeply disappointing to recognize that we know relatively little about the net impact of most of these social programs“ (Savedoff et al., 2006, S. 1).

Als zentrale Ursachen für den noch immer vorherrschenden „[...] Mangel an ‚echten‘ Wirkungsevaluationen“ (Borrmann & Stockmann, 2009a, S. 49) und der daraus resultierenden Evaluationslücke gelten die *Quantität* („zu wenige“) und *Qualität* („methodisch unzureichend“) von Wirkungsevaluationen.¹⁹⁴ Die Arbeitsgruppe appellierte an Entwicklungsländer und EZ-Organisationen, die Evaluationslücke durch die Nutzung bewährter Verfahren zu schließen und eine unabhängige Einheit zu gründen, durch die Ansätze und Methoden (weiter-)entwickelt sowie *rigorose Wirkungsevaluationen* finanziert und durchgeführt werden können. Der Appell führte zur Gründung der internationalen Initiativen *Network of Networks on Impact Evaluation (NoNIE)* und der *International Initiative on Impact Evaluation (3IE)* (vgl. Borrmann & Stockmann, 2009a, S. 50; Caspari, 2008, S. 138ff; Caspari, 2009, S. 189f; Caspari & Barbu, 2008, S. 4; OECD, 2013, S. 37).¹⁹⁵

In Deutschland veranlasste das BMZ Ende der 1990er Jahre eine erste sog. *Systemprüfung* der Evaluierung in der dt. EZ (vgl. Borrmann et al., 1999)¹⁹⁶, die zweite erfolgte im

¹⁹³ Das CGD ist „[...] ein gemeinnütziger ‚Think Tank‘ mit Sitz in Washington D.C.“ (Caspari & Barbu, 2008, S. 2).

¹⁹⁴ Zum einen werden zwar hohe Summen in M&E-Aktivitäten gesteckt, jedoch nicht in umfassende Wirkungsevaluationen. Zum anderen sind durchgeführte Wirkungsevaluationen methodisch unzureichend, da selten rigorose Methoden genutzt werden (Caspari & Barbu, 2008).

¹⁹⁵ Weitere Ausführungen zu den Netzwerken s. unter *Institutionalisierung und Professionalisierung*.

¹⁹⁶ Immerhin wiesen zur Zeit der ersten Analyse alle untersuchten EZ-Organisationen Evaluationsstrukturen auf „[...] in rudimentary form at the very least“ (Borrmann, 1999, S. 256).

Jahr 2000 (vgl. Borrmann, Gleich, Holthus & Shams, 2001)¹⁹⁷ und eine dritte wurde 2007/2008 durchgeführt (vgl. Borrmann & Stockmann, 2009a, 2009b)¹⁹⁸. Hierbei wurden „[...] die Evaluationssysteme des BMZ und der staatlichen und nichtstaatlichen EZ-Institutionen im Hinblick auf ihre Organisation, Konzeption, Methoden und Verfahren sowie Umfang und Struktur [...]“ bewertet, Ansatzpunkte zur Verbesserung gegeben sowie ab der zweiten Untersuchung Veränderungen und Fortschritte analysiert (Borrmann & Stockmann, 2009a, S. 31). Die Ergebnisse jeder Studie zeigten, dass die EZ-Organisationen zwar maßgebliche Anstrengungen unternommen hatten, um ihre Evaluationssysteme weiterzuentwickeln, allerdings hatten Defizite bzgl. dreier Kernbereiche seit der ersten Studie Bestand: *Unabhängigkeit und Objektivität* von Evaluationen, *ex-post Evaluationen mit dem Fokus auf Wirkungen und Nachhaltigkeit* und der *Einbindung von Partnern (partizipative Ansätze)* (Borrmann, 2009). Zwar schnitt das deutsche EZ-Evaluationssystem im OECD-Ländervergleich „recht gut ab“ (Borrmann & Stockmann, 2009a, ebd., S. 32), allerdings wurde i.R. der dritten Untersuchung¹⁹⁹ der Zustand des Systems, welches „[...] unter enormer institutioneller, konzeptioneller, terminologischer und methodischer Heterogenität [leidet]“, beklagt (ebd., S.17), was auf die generell stark fragmentierte deutschen EZ-Struktur zurückzuführen war:²⁰⁰ „Im Grunde gibt es so viele Evaluationssysteme wie es EZ-Organisationen gibt, und keines gleicht dem anderen“ (Stockmann & Borrmann, 2009, S. 291). Dem BMZ wird hierbei als federführendem Ministerium eine *mangelnde Steuerungsmacht* diagnostiziert (ebd., S. 292). Zur Überwindung

¹⁹⁷ Die Ergebnisse der zweiten Untersuchung zeigten positive Bemühungen, aber auch, dass „[...] der eingeschlagene Reformkurs keineswegs abgeschlossen war“ (Borrmann & Stockmann, 2009a, S. 31).

¹⁹⁸ Hier konnten zahlreiche positive Veränderungen auch als Ergebnis nationaler Reformbemühungen und vor dem Hintergrund sich verändernder EZ-Ziele und Instrumente nachgewiesen werden (bspw. Aufstockung des Evaluationspersonals, eine klare Ausrichtung auf Wirkungen, v.a. bei der GTZ) (Stockmann & Borrmann, 2009).

¹⁹⁹ Neben dem BMZ wurden 19 deutsche EZ-Organisationen untersucht, die EZ-Vorhaben ganz oder teilweise im Auftrag und mit Mitteln des BMZ durchführen: Sieben staatliche DO (u.a. KfW, GTZ, InWEnt, DED), vier kirchliche Organisationen (bspw. Misereor, Caritas), sechs politische Stiftungen (bspw. Friedrich-Ebert- und Heinrich-Böll-Stiftung) und zwei NGOs (bspw. Deutsche Welthungerhilfe) (vgl. Stockmann & Borrmann, 2009, S. 270).

²⁰⁰ „Diese Vielzahl geberseitiger Akteure und die starke institutionelle Trennung von finanzieller und technischer EZ haben bei anderen Gebern kaum Parallelen und sind seit Langem Gegenstand nationaler und internationaler Kritik“ (Stockmann & Borrmann, 2009, S. 273).

der Herausforderungen wurden i.R. der dritten Studie drei konkrete Szenarien vorgestellt²⁰¹, die neben dem internationalen Druck letztendlich zur *Zusammenführung aller Organisationen der technischen Zusammenarbeit*, d.h. GTZ, InWent (Internationale Weiterbildung und Entwicklung) und Deutscher Entwicklungsdienst (DED), im Jahr 2010²⁰² zur GIZ sowie zur *Gründung eines unabhängigen Evaluierungsinstituts* im Jahr 2012 führte.²⁰³ Die befragten Organisationen nutzten Evaluationen vorrangig zur *Entscheidungsfindung* und *Steuerung* von Projekten und für das institutionelle Lernen (*Entwicklung*). Die Funktionen der Rechenschaftslegung und Außendarstellung (*Legitimierung*) wurden weniger und mit einer niedrigeren Priorität genannt. „Der *Kontrollaspekt* von Evaluationen, der jahrzehntelang zumindest in der deutschen staatlichen EZ-Evaluation im Vordergrund stand, ist mittlerweile offensichtlich so in den Hintergrund geraten, dass er von keiner Organisation auch nur erwähnt wurde“ (Borrmann & Stockmann, 2009a, S. 130).

2010 wurde die bis dato größte *Review* des DAC mit dem Titel „Evaluation in Development Agencies“ (OECD, 2010c) veröffentlicht.²⁰⁴ Sie identifizierte wichtige Trends und aktuelle Herausforderungen bei der EZ-Evaluation von 32 Mitgliedern.²⁰⁵ Im Vergleich zu früheren Erhebungen wurde bei dieser Studie ein deutlicher Anstieg an Evaluationen registriert. Durchschnittlich führten die Organisationen 24 Evaluationen pro Jahr durch, bilaterale Organisationen gaben hierbei im Schnitt 2,4 Mio. USD aus. Deutschland liegt mit 89 Evaluationen pro Jahr (BMZ, GTZ und KfW) und mit Ausgaben von 7,38 Mio. USD pro Jahr deutlich über dem Durchschnitt. Setzt man die Ausgaben allerdings in Relation

²⁰¹ A: Ausbau der Evaluationseinheit des BMZ; B: Gründung einer unabhängigen Evaluationsagentur; C: Gründung eines unabhängigen Evaluationsbeirats (s. Stockmann & Borrmann, 2009, S. 288ff).

²⁰² Scholz und Wolff (2010) beleuchten die institutionellen Reformen, inklusive involvierter Studien und Gutachten sowie daraus resultierende Empfehlungen und Umsetzungen ausführlich (vgl. ebd., S. 240ff).

²⁰³ S. hierzu Ausführungen unter der Überschrift *Institutionalisierung und Professionalisierung*.

²⁰⁴ Ashoff (2013) gibt einen Überblick zu den *Peer Reviews*, reflektiert diese kritisch, betrachtet Veränderungen, ihre Bedeutung und Zukunft. „Laut DAC-Sekretariat wurden in den letzten beiden Jahren [2011 und 2012] 88% aller Empfehlungen ganz oder teilweise umgesetzt“ (ebd., S. 1).

²⁰⁵ Evaluationsaspekte sind: Institutionelle Rahmenbedingungen (Strukturen, Richtlinien, Praktiken), Ressourcen, Unabhängigkeit der Evaluationsfunktion, Berichterstattung und Nutzung von Ergebnissen, Managementsysteme, Geberkoordinierung, gemeinsame Evaluationen, Einbeziehung der Partnerländer (OECD, 2013, S. 36). Ergebnisse der Studie werden z.T. im folgenden Text verortet.

zu den gesamten öffentlichen EZ-Ausgaben, was im Falle Deutschlands 0,12% ausmacht, befindet sich Deutschland im Mittelfeld (vgl. ebd., S. 25).²⁰⁶

Die Studie zur Implementierung der Paris-Deklaration (Wood et al., 2011), die auch eine *Meta-Evaluation* beinhaltet (Caspari, 2015), wird seitens der OECD (2013, S. 21) als wichtigste und als „the largest evaluation ever undertaken“ bezeichnet.²⁰⁷ Die Ergebnisse, die 2011 in Busan präsentiert wurden, zeigen, dass die Umsetzung der Paris-Erklärung langsam und unregelmäßig verlief. Die Gebenden haben hierbei weniger Fortschritte gemacht als die Partnerländer (Ashoff, 2011). Die Evaluation wurde 2012 mit dem *Outstanding Evaluation Award* der *American Evaluation Association* ausgezeichnet (OECD, 2013).

Schwerpunkte und Ansätze

Die Unvereinbarkeit der Evaluationsergebnisse auf *Mikro- und Makro-Ebene* blieb weiterhin Gesprächsgegenstand: In Evaluationsberichten unterschiedlicher entwicklungs-politischer Organisationen wurden „[...] mehr als drei Viertel der durchgeführten Projekt- oder Programminterventionen als zufriedenstellend oder besser bewertet“, wohingegen statistische Ländervergleiche²⁰⁸ auf der Datengrundlage der letzten Dekaden (1970-2005) „keine robusten Ergebnisse“ liefern konnten (Faust, 2010, S. 41).²⁰⁹

Laut Roxin (2016) fand die *methodische Debatte deutlich differenzierter* statt. Nachdem zur Jahrtausendwende vermehrt *ex-post-Evaluationen* auf Projektebene durchgeführt wurden, kamen im Zuge der Entwicklung des *Poverty Social Impact Assessment (PSIA)*

²⁰⁶ Im internationalen Vergleich rangiert Deutschland in der Bewertung der allgemeinen Evaluationskultur ebenfalls nur im Mittelfeld (vgl. Furubo et al. 2002; Jacob, Speer & Furubo, 2015).

²⁰⁷ Durch die Unterzeichnung verpflichteten sich die Partner*innen dazu, ihre M&E Prozesse untersuchen zu lassen, um verstehen zu können, „how increased effectiveness contributes to meeting development objectives“ (OECD, 2005/2008, S. 3). Insgesamt nahmen ab 2007 22 Partnerländer und 18 Geberorganisationen an der Evaluation teil (OECD, 2013; vgl. Wood et al., 2011). Ashoff (2011) fasst zentrale Ergebnisse zusammen.

²⁰⁸ Hier werden ökonomische und soziale Indikatoren (bspw. Wirtschaftswachstum, durchschnittliches Pro-Kopf-Einkommen, Kindersterblichkeitsrate) als zu erklärende Variablen genutzt, „[...] ob der nach Ländern variierende Zufluss von öffentlichen EZ-Transfers [...]“ einen positiven Einfluss hatte (Faust, 2010, S. 42). Schlüsselbegriffe und methodische Herausforderungen erläutern Faust und Leiderer (2010).

²⁰⁹ Nach Faust (2010, S. 42) lässt sich das Paradoxon erklären, laut Neubert (2010, S. 139) existiert es nicht, sondern ist „[...] alleine der unzureichenden Untersuchung der Phänomene von beiden Seiten her geschuldet“. Erläuterungen s. Neubert (2010, S. 149ff.), Ashoff (2010, S. 46), Faust (2010, S. 42), Hemmer (2010, S. 221ff.).

der Weltbank um 2003 *ex-ante Untersuchungen* in Mode. Hierbei wurde laut Neubert (2010, S. 143) ein „wichtiger Schritt vollzogen“, da erstmalig mögliche negative Wirkungen und Risiken systematisch von Beginn an in Betrachtung gezogen wurden. In Deutschland herrschten in der GTZ Selbstevaluationen (*intern*) und *ex-ante* Evaluationen vor, während die KfW verstärkt *ex-post*-Evaluationen durchführte (Neubert et al., 2014).²¹⁰

Durch den neuen Auftragsrahmen mit dem BMZ von 2002 verpflichtete sich die GTZ, „[...] ihre unternehmerischen Erfolge an der Wirksamkeit der EZ-Maßnahmen messen zu lassen. Hierfür hat die GTZ ihr gesamtes Auftragsmanagement, das M&E-System, das Berichtswesen sowie das Evaluierungssystem auf Wirkungen ausgerichtet“ (Stockmann & Borrmann, 2009, S. 278).²¹¹ Das sog. *Wirkungsorientierte Monitoring (WoM)* „ist kein grundlegendes neues Instrument, sondern bezeichnet lediglich eine *Form des Monitoring*, das an Wirkungen orientiert ist [...]“ (Oberndörfer et al., 2010).²¹² WoM ist „integraler Bestandteil des Programmmanagements“, wobei „[...] der gesamte Veränderungsprozess beobachtet [wird], der durch ein Programm ausgelöst wird“ (Schweitzer, 2013, S. 6).

Basierend auf der verbreiteten Einführung von theoriegeleiteten RBM-Systemen, wodurch Daten kontinuierlich gesammelt und Projekte in Programmeinheiten integriert werden können, setzte sich der rückläufige Trend an Projektevaluationen hin zu vermehrt strategischeren thematischen²¹³, sektorübergreifenden oder landesweiten *Evaluationen und Evaluationssynthesen* fort (vgl. OECD, 2010, S. 21). Querschnittsthemen wie Gender und Umwelt werden 2009 mittlerweile routinemäßig in allen Evaluationen in einer Reihe von Agenturen berücksichtigt (ebd.). Seit 2013 hat sich in der *European*

²¹⁰ Beide Organisationen führen im Auftrag des BMZs Einzelevaluationen nach der Beendigung ihrer Maßnahmen durch: Die KfW flächendeckend von 1988 bis 2006 und auf Stichprobenbasis seit 2007, die GTZ auf Stichprobenbasis seit 2005. Die KfW nutzt hierbei ein System der Notenvergebung und führt diese Schlussprüfung ca. 3-5 Jahre nach Beendigung der Maßnahme durch (Hemmer, 2010). Für weitere Details s. Hemmer (2010, S. 209ff.), Terberger (2011, S. 228ff.) und Zintl (2009, S. 246ff.).

²¹¹ Diese Neuausrichtung steht „[...] in einem starken Kontrast zu dem tatsächlichen Wissen über die verursachten Wirkungen [...]“, was vor allem damit zusammenhängt, dass die Evaluationspraxis kaum Veränderungen erfahren hat (bspw. hinsichtlich der Evaluationsdesigns und Methoden) (ebd., S. 279).

²¹² Hervorhebungen durch die Autorin.

²¹³ Auch hierzu entwickelte die Expert*innengruppe Leitfäden, z.B. „Evaluating Humanitarian Assistance in Complex Emergencies“, „Evaluating Peace Building Activities“ (s. OECD, 2013, S. 34).

Evaluation Society (EES) eine thematische Arbeitsgruppe „Evaluating Sustainable Development“ etabliert (Martinuzzi & Meyer, 2017).

Die Rufe nach mehr, besserer und kausaler Wirkungsevaluationen (u.a. Savedoff et al., 2006; Duflo²¹⁴, Glennerster & Kremer, 2007) löste einen Trend zu *rigorosen Evaluationsdesigns* (experimentelle und quasi-experimentelle Verfahren) aus, wie bspw. dem randomisierten Experiment (RCT), bei dem Organisationen sich bereits vor der Maßnahme mit der Analyse von Wirkungen beschäftigen müssen, d.h. Evaluation bleibt nicht länger als eine „lästige, isolierte Pflichtübung am Projektende“ übrig (Faust, 2010, S. 45; vgl. auch Caspari, 2008, S. 190ff.; Caspari & Barbu, 2008, S. 5ff.; Neubert, 2010, S. 146f., Leiderer, 2010, S. 241ff; White & Barbu, 2006, S. 2ff.).²¹⁵ Die Ergebnisse rigoroser Wirkungsevaluationen werden zunehmend i.R. *systematischer Reviews* genutzt und aggregiert, wodurch die „Validität der Aussagen erhöht“ wird (Bruder, Faust & Krämer, 2019, S. 3). In sog. *Karten der Evidenzlücken (evidence gap maps)* wird die Evidenz in einzelnen Themenfeldern abgebildet (ebd.). Insbesondere die *International Initiative for Impact Evaluation (3iE)* hat sich auf die Bereitstellung solcher Landkarten spezialisiert (vgl. 3iE, 2021).²¹⁶

Hinsichtlich *qualitativer Designs und partizipativer Ansätze* zur Wirkungsevaluation wurden in den Nullerjahren zahlreiche Verfahren entwickelt, darunter das computerbasierte Interviewsystem zur Selbstevaluierung *e-VAL* der GTZ (s. Reuber, 2006, S. 81), um Schwächen wie „die geringe Systematisierung und Aggregierbarkeit der Ergebnisse“ oder der selektive Charakter auszugleichen (Neubert, 2010, S. 144). Sie zeichnen sich u.a. durch ihre übersektorale Anwendbarkeit, die Einbeziehung ungeplanter Wirkungen sowie Vergleichsmöglichkeiten durch Quantifizierung aus. Aufgrund der Vielzahl an Ansätzen und einem Dokumentationsmanko veröffentlichte die DeGEval eine *Landkarte über Verfahren der Wirkungsanalyse* (Oberndörfer et al., 2010), die alle praxisbewährten

²¹⁴ Die Ökonomin *Esther Duflo* wird von Neubert et al. (2014, S. 65) als „Galionsfigur“ der experimentellen Designs bezeichnet, da sie „diese Verfahren für die EZ zum Durchbruch führte“. Für ihre Forschungsarbeit mit dem Schwerpunkt auf der Wirksamkeit von EZ-Maßnahmen zur Armutsbekämpfung erhielt Duflo 2019 den Wirtschaftsnobelpreis (Groll, 2019, 14. Oktober). S. auch *Schwerpunkte, Ansätze* unter Kap. 4.5.

²¹⁵ Solche umfangreichen quantitativen Designs fordern sowohl personelle Kapazitäten als auch finanzielle Ressourcen, beidem mangelte es laut der OECD-Studie (vgl. OECD, 2010, S. 21, 26) zahlreichen Evaluationseinheiten auch Ende der Nullerjahre.

²¹⁶ S. Ausführungen unter *Professionalisierung und Institutionalisierung*.

qualitativen und quantitativen Ansätze darstellt und ihre Eignung und Charakteristika beleuchtet.²¹⁷ Um Stärken und Schwächen der jeweiligen Designs auszugleichen, ging die EZ zu einem durchdachten *quantitativ-qualitativen Methodenmix* über, wobei der Evaluationsgegenstand die Nutzung angepasster Methoden nach sich zieht (Neubert et al., 2014; Roxin, 2016).

Komplexe Zielsysteme wie *Armut, Good Governance* usw. oder der Zusammenhang zwischen EZ-Geldern und der Entwicklung von Wirtschaftskennzahlen werden immer häufiger mittels *quantitativer Indikatoren* gemessen. Die *Aggregatdatenforschung*, die auf großen Datensets basiert, wird zunehmend wichtiger, sie standardisiert und globalisiert EZ-Programme (Neubert et al., 2014). Die Ergebnisse dieser Wirkungsanalysen auf der Basis quantitativer Makrodaten sind allerdings oft wenig robust (vgl. Faust & Leiderer, 2010, S. 166ff) und ihre Aussage begrenzt (vgl. Nuscheler, 2008, S. 25), u.a. auch aufgrund mangelnder Datengrundlagen (Neubert et al., 2014).

Durch die zunehmende Komplexität der EZ-Programme breitete sich im Verlauf der Nullerjahre ein *systemisches Verständnis von Evaluation* aus. Die GTZ entwickelte in den frühen 2000er Jahren ein noch lineares, monokausales *Wirkungskettenmodell*, das im Zuge der institutionellen Fusion durch ein multikausales und vernetztes *Wirkungsgefüge* abgelöst wurde und bis heute genutzt wird (Neubert et al., 2014).²¹⁸ Durch die Abwendung von linearen Modellen hin zu systemischen können jetzt verschiedene Ebenen von Wirkungen erfasst werden, d.h. dass auch nicht-intendierte sowie Wirkungsbeiträge anderer Akteur*innen und unterschiedliche Erhebungsmethoden zur Erfassung von Wirkungen zum Einsatz kommen (Roxin, 2016).

Im Hinblick auf die Schwerpunkte und der Konstitution von Studien setzte sich international der Trend der gemeinsamen Evaluationsaktivitäten (*Joint Evaluations*) fort (OECD,

²¹⁷ U.a. Outcome Mapping, PIM, WoM und der CEval-/ Stockmann'sche Ansatz (Oberndörfer et al., 2010, S. 35).

²¹⁸ Neubert (2010, S. 143) und Neubert et al., (2014, S. 63) äußern sich zu den linearen Modellen kritisch und sind der Ansicht, dass die Fokussierung auf lediglich intendierte Wirkungen einen „Tunnelblick“ auslöse und die Gefahr bestünde, dass „Wunschbilder“ konstruiert werden, ohne dass eine weitere Überprüfung stattfindet. „Insbesondere die Evaluierung von Nachhaltigkeit kam bei den Evaluierungen oft zu kurz [...]“ (ebd., S. 63).

2013),²¹⁹ woran sich auch Deutschland beteiligte (vgl. Zintl, 2006, S. 49). Aufgrund hoher Transaktionskosten und des Drucks den ‚eigenen‘ EZ-Beitrag eines Landes darzustellen, gab es in den letzten Jahren allerdings einen deutlichen Rückgang. In Deutschland führte das BMZ in den 2000ern vorrangig „[...] komplexere und ‚politischere‘ Evaluationen durch“, u.a. *Evaluationssynthesen* (Querschnittsauswertung und Synthese) von Vorhaben in einzelnen Sektoren, Programmevaluationen auf Länderebene²²⁰ oder Evaluationen von Instrumenten der EZ (z.B. personelle Zusammenarbeit, Sozialstrukturförderung) (Bliss, 2006, S. 15). *Meta-Evaluationen* sind international und auch in der dt. EZ rar. Die GTZ/GIZ und InWEnt geben bis 2014 insgesamt sechs Meta-Evaluationen in Auftrag, was Caspari (2015, S. 155) als „vergleichsweise positiv hervorstechen[d]“ bezeichnet.

Mitte der 2000er wurde die *Evaluation von multilateralen Organisationen* seitens der DAC-Expert*innengruppe mittels eines *Peer-Review-Ansatzes* erneut angegangen, getrieben von der Hoffnung, zuverlässige Evaluationssysteme zu identifizieren, die von anderen Stakeholdern genutzt werden können.²²¹ Ab 2004 arbeitete die Expert*innen-gruppe systematischer an der Stärkung von Evaluationskapazitäten der OECD-Mitglieder. Hierfür wurde ein *Assessment Tool for DAC Peer Reviews and Evaluation Managers* entwickelt, das i.R. zahlreicher Studien genutzt wurde und auch zur *Selbstevaluierung* diente (OECD, 2013).

Mayring et al. (2017) führten basierend auf den Beiträgen der *Zeitschrift für Evaluation* eine *Review* nach Evaluationsthemen von 2002–2016 durch. Die Ergebnisse zeigen, dass die *Methodenwahl* in Deutschland im Vordergrund des Evaluationsinteresses stand, es in dem Themenbereich also eine intensive Auseinandersetzung über einen langen Zeitraum gab. Ab 2010 gewinnen die Themen *Wirkungsnachweis* und *Professionalisierung* an Relevanz (ebd.). Letzteres wird im folgenden Kapitel beleuchtet.

²¹⁹ Ab Ende der 1990er wurden zahlreiche Joint Evaluations seitens der DAC-Expert*innengruppe durchgeführt und Leitlinien veröffentlicht. S. hierzu Auflistungen in OECD (2013, S. 17; 20; 22f.; 34).

²²⁰ Wobei das gesamte Länderportfolio der deutschen staatlichen EZ (GTZ, KfW, DED, InWEnt, Centrum für internationale Migration und Entwicklung (CIM)) i.R. festgelegter Schwerpunktbereiche (z.B. Bildung, Energie oder Ressourcenschutz) analysiert wurde (vgl. Reuber, 2006, S. 80).

²²¹ 2005 das UNDP und 2006 UNICEF. Die methodische Vorgehensweise des Ansatzes wurde im Anschluss verfeinert. Zahlreiche weitere multilaterale Organisationen wurden danach mit dem entwickelten Ansatz begutachtet. S. hierzu die Auflistung in OECD (2013, S. 25).

Institutionalisierung und Professionalisierung

Die Internationalisierung von EZ-Ansätzen, u.a. durch die MDGs, brachte eine Internationalisierung von EZ-Evaluierung mit sich und damit auch eine Institutionalisierung und Professionalisierung auf internationaler Ebene. Als Indikatoren für eine *zunehmende Professionalisierung* im Bereich der Evaluation sieht Stockmann (2018) neben der steigenden Zahl an Fachzeitschriften, der Etablierung von Evaluationsstandards, der weltweit angebotenen Weiterbildungen²²² auch die *Zunahme an nationalen und internationalen Evaluationsgesellschaften und Initiativen*. Wie bereits erwähnt, gab es Mitte der 1990er Jahre nur wenige Evaluationsvereinigungen, -gesellschaften, -verbände oder -netzwerke, die übergeordnet als *Voluntary Organizations for Professional Development (VOPEs)* bezeichnet werden. Zu Beginn der 2000er erhöhte sich die Zahl auf über 30 registrierte Organisationen, Ende 2013 wurden bereits mehr als 180 gezählt (Rodríguez-Bilella & Lucero, 2017; IOCE, 2019). An der Gründung der neuen VOPEs waren immer multilaterale EZ-Organisationen (UN, Weltbank) oder bereits bestehende Evaluationsnetzwerke beteiligt. Für den Bereich der EZ-Evaluation spielen v.a. die *International Development Evaluation Association (IDEAS)*, die *International Organisation for Cooperation and Evaluation (IOCE)*²²³, die Vereinigung bestehender Netzwerke *Network of Networks on Impact Evaluation (NoNIE)*, die unabhängige Organisation *International Initiative for Impact Evaluation (3iE)*²²⁴ und *EvalPartners* eine wichtige Rolle (Rodríguez-Bilella & Lucero, 2017).²²⁵ Besonders die Gründung von NONIE und 3iE sehen Expert*innen vor dem Hintergrund der Wirksamkeitsdebatte von besonderem Mehrwert (Faust, 2010; Neubert, 2010; Neubert et al., 2014; Oberndörfer et al. 2010).²²⁶ Durch die Etablierung

²²² Meyer und Stockmann (2016) ermittelten 13 solcher Fachzeitschriften, vier davon in Europa (Deutschland, Schweiz, Spanien, Großbritannien) sowie 24 Masterstudiengänge weltweit im Bereich der Evaluation.

²²³ Seit Ende der 1990er Jahre bemühen sich „[...] Mitglieder der meisten größeren nationalen Evaluationsgesellschaften um die Schaffung einer weltumspannenden Mittlerorganisation, die den Austausch von Wissen, Konzepten und Erfahrungen mit der Evaluation interkulturell fördern soll“, was zur Gründung der internationalen Organisation führte (Mayring et al., 2017, S. 24).

²²⁴ 3iE finanziert sich aus der Zuwendung verschiedener Staaten und aus privaten Stiftungen wie etwa der Gates-Stiftung (Faust, 2010; Levine & Savedoff, 2007, 1. Dezember).

²²⁵ Rodríguez-Bilella und Lucero (2017, S. 56ff.) beschreiben die VOPEs im Detail.

²²⁶ 2009 wurde die „NoNIE Guidance on Impact Evaluations“ veröffentlicht. Im selben Jahr organisierten UNICEF, 3iE und NoNIE eine Konferenz in Kairo mit dem Titel „Perspectives in Impact Evaluation“, die auch von Evaluierenden der Partnerländer besucht wurden (OECD, 2013). Weitere Details s. Caspari und Barbu (2008), Leeuw und Vaessen (2009), White & Bamberger (2008). Die NoNIE-Homepage existiert nicht mehr (Stand: 19.02.2022).

dieser beiden Plattformen konnten zahlreiche Beispiele zur Wirkungsevaluation gesammelt werden, „[...] die wiederum weltweit Lerneffekte für die Ausgestaltung zukünftiger Programmansätze ermöglichten“ (Neubert et al., 2014, S. 65).²²⁷

Auch auf europäischer Ebene standen ab den 2000ern zahlreiche Reformen hinsichtlich der gemeinschaftlichen EZ auf der Agenda.²²⁸ Die Europäische Union (EU), die seit 1993 eine eigenständige Kompetenz für EZ hat, gründete 2001 *EuropeAid*, um die Planung, Abwicklung und Evaluation von europäischen Maßnahmen zu vereinheitlichen (Grimm, 2010).²²⁹

Die Kritik an der mangelnden Unabhängigkeit von Evaluation hat in zahlreichen EZ-Agenturen zur *Etablierung interner unabhängiger Abteilungen oder Einheiten* geführt. In Deutschland gründete die KfW bereits 1999 eine unabhängige Evaluierungseinheit, die 2000 ihre Arbeit aufnahm (Jahn, 2006). Die GTZ richtete Anfang der Nullerjahre die *Stabsstelle für Evaluierung* ein, die von den operativen Bereichen getrennt ist (Reuber, 2006; Borrmann & Stockmann, 2009b).²³⁰ Ab 2005 beauftragte die Stabsstelle auch *Fremdevaluierungen (extern)*, die durch Beratungsfirmen oder andere Auftragsnehmende durchgeführt werden (Reuber, 2006). Diese externen Gutachtenden werden u.a. auch mit Wirkungsevaluationen von Maßnahmen beauftragt. Da dies regelmäßig im Auftrag der Durchführenden geschieht, bleibt die Kritik am ‚positiven Bias‘ und stellt die Unabhängigkeit in Frage (Roxin, 2016).

Wegen der bereits erwähnten nationalen und internationalen Kritik am fragmentierten deutschen EZ-System und dem daraus resultierenden defizitären Evaluationssystem gründete das BMZ 2005 mit den Durchführungsorganisationen GTZ, InWEnt, KfW und

²²⁷ I.R. einer seitens des BMZ beauftragten Studie (vgl. Caspari & Barbu, 2008) wurden 29 Wirkungsevaluationen, veröffentlicht auf der NONIE-Homepage, analysiert. Die Meta-Evaluation zeigte, dass auch in der internationalen Wirkungsevaluation noch Schwächen zu überwinden sind. Deutschland war hierbei mit keiner Studie vertreten (Borrmann & Stockmann, 2009a).

²²⁸ Die Entwicklungspolitik der EU ist kompliziert organisiert. Grimm (2010) beschreibt Strukturen und Instrumente im Detail. Die EU ist weltweit der größte Geber öffentlicher Entwicklungsgelder, da mehr als die Hälfte der öffentlichen EZ-Ausgaben aller DAC-Mitglieder (57%) Gelder aus der EU sind (ca. 20% in Dtl.) (ebd.).

²²⁹ *EuropeAid* erstellt Jahresberichte zur EU-Außenhilfe. Die Europäische Kommission veröffentlicht in ihren Jahresberichten aggregierte Projektdaten (Grimm, 2010). Die *EU Evaluation Policy*, in der Evaluation als „key learning tool“ bezeichnet wird (EU, 2014, 14. August, S. 3), wurde 2014 online veröffentlicht.

²³⁰ Borrmann und Stockmann (2009b, S. 93ff.) schildern ausführlich die Aufgaben der Stabsstelle.

DED die *Arbeitsgruppe Evaluierung aus einem Guss*. Doch die angestrebte *Evaluierung aus einem Guss* scheiterte an extrem zeitaufwendigen und mühsamen Prozessen und den mangelnden Steuerungskapazitäten des BMZ (vgl. Borrmann & Stockmann, 2009a, S. 25; Stockmann & Meyer, 2017a, S. 85). Letztendlich wurde im Jahr 2012 das *Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval)* gegründet.²³¹ Das Institut orientiert sich an sieben Prinzipien: unabhängig, bedarfsorientiert, qualitätsbewusst, partnerschaftlich, partizipativ, effizient in den Evaluierungen und transparent in den Ergebnissen (Roxin & Asche, 2013). Handlungsfelder sind (1) strategisch relevante Evaluierungen, (2) Weiterentwicklung von Methoden und Standards, (3) Förderung von Evaluierungskapazitäten, (4) Wissensvermittlung und Vernetzung (DEval, 2019a). Durch die Etablierung wurden mehr *Reviews* allgemein sowie *Evaluationssynthesen*, *Meta-Evaluationen* (bspw. Evaluationen zu bestimmten Sektoren, EZ-Instrumenten) und *Wirkungsevaluationen* im Speziellen möglich, die eine Gesamtsicht der dt. EZ, komplexe Zusammenhänge und politisch relevante Fragestellungen fassen (Neubert et al., 2014; Roxin, 2016).²³² Aus Sicht des BMZ bildet die Gründung des DEval „[...]den Schlussstein der Strukturreform mit der Fusion der Vorfeldorganisationen der technischen Zusammenarbeit der vergangenen Jahre“ (Roxin, 2016, S. 16).

Im Hinblick auf die Aspekte *Transparenz* und *Rechenschaftslegung* bemühte sich die DAC-Expert*innengruppe zur Jahrtausendwende, Evaluationsergebnisse online zur Verfügung zu stellen. Vorerst konnte der Dienst nur intern genutzt werden, ab 2005 wurde dann eine öffentlich zugängliche Webseite (*DAC Evaluation Resource Cente*) ins Leben gerufen, die damals bereits 1300 Evaluationsberichte enthielt.²³³ Ebenfalls online werden seit 2011 Kurzdossiers unter dem Titel „Evaluation Insights“ veröffentlicht, in denen Evaluationsergebnisse zu aktuellen Debatten hervorgehoben werden (OECD, 2013). Auch das BMZ und die staatlichen Durchführungsorganisationen stellten ab Mitte der 2000er vermehrt Berichte oder zumindest Zusammenfassungen online zur Verfügung (Bliss, 2006). Doch trotz der Etablierung unabhängiger Stabsstellen, der Neugründung

²³¹ Von der Idee eines unabhängigen Evaluationsinstituts bis zur Institutionalisierung war es ein langer Weg. S. hierzu Stockmann (2012).

²³² Solche Evaluationen waren bis dato im BMZ-Referat für Erfolgskontrolle verortet (Neubert et al., 2014).

²³³ Seitdem wurde die Webseite um Funktionen und Links ergänzt, Dokumente und Berichte wurden aktualisiert sowie Facebook, Twitter und YouTube verwendet, um Nachrichten an eine breitere Öffentlichkeit zu übermitteln (OECD, 2013, S. 40).

von Institutionen sowie der Bemühungen um die Offenlegung von Evaluationsergebnissen werden zu wenige Evaluationsergebnisse genutzt und/oder verbreitet (vgl. OECD, 2010c, S. 10). Die Veröffentlichungspraxis der EZ-Organisationen war und ist zahlreicher Kritik ausgesetzt (vgl. OECD, 2010c). Die Zurückhaltung wird als „Vermeidung von Imageverlust“ bezeichnet (Caspari, 2004, S. 27), der Evaluation die Funktion von „Außendarstellung“ zugeschrieben und Ergebnisse nach wie vor in Form von „Hochglanzbroschüren“ veröffentlicht, die „annähernd unkritisch [...] vornehmlich die Erfolge der eigenen Arbeit“ veranschaulichen (vgl. Borrmann et al., 1999, S. 17, 26).

4.5 Phase V: Alte und neue Herausforderungen (seit 2015)

„Evaluation everywhere, at every level, will play a key role in implementing the development agenda.“ (UNEG, 2015, 25. September)

Dass Evaluation Teil der *Agenda 2030* ist und explizit erwähnt wird, ist ein Meilenstein für die globale Evaluationsgemeinschaft und setzt sie zeitgleich unter großen Druck.²³⁴ „*Evidenzbasierte Politik* ist das Schlagwort, das die Messlatte für Politik und Praxis, aber vor allem auch für die Evaluierung höher legt“ (Hauck, 2019, S. 134). Die in der Agenda implizierten „Annahmen“, z.B. national erhobene Daten seien aktuell und ausreichend belastbar, um die Erreichung einzelner SDG-Indikatoren darzustellen oder alle Länder verfügten über ausreichende Evaluationskapazitäten, sehen in der Realität anders aus (vgl. Frueh, 2019, S. 28).

Evaluation nimmt eine „[...] Schlüsselfunktion für die Umsetzung nachhaltiger Entwicklung ein“ (DEval, 2019, S. 99). Um der *Komplexität* und *Mehrdimensionalität* des Zielsystems gerecht zu werden, bedarf es einer *Weiterentwicklung der Evaluation*, bspw. durch die Erweiterung ihrer Funktionen, die Nutzung neuer Datenmengen und -quellen, einer verstärkten interdisziplinären Zusammenarbeit und die Nutzung einer Methodenvielfalt (ebd.).

²³⁴ Evaluation ist Teil des Artikels 74 „follow-up and review processes“, s. UN General Assembly (2015b, S. 32f.).

Kontext und Meilensteine

Mit dem Beschluss zur *Agenda 2030* für nachhaltige Entwicklung und der beinhalteten Weiterentwicklung der *MDGs* zu *17 Sustainable Development Goals (SDGs)* im Jahr 2015 „[...] hat sich die internationale Gemeinschaft einen ambitionierten Referenzrahmen geschaffen, der insbesondere für die Entwicklungspolitik und EZ von großer Bedeutung ist“ (Faust, 2019, S. 7). Die *Mehrdimensionalität* des Zielsystems, d.h. die umfassende Beschreibung nachhaltiger Entwicklung, die nicht nur die Dauerhaftigkeit von Wirkungen meint, sondern auch die sozialen, ökonomischen und ökologischen Wechselwirkungen umfasst, und die damit verbundene *Komplexität* (17 SDGs, 169 Unterziele und dazugehörige 232 Indikatoren sowie Prinzipien ihrer Umsetzung)²³⁵ stellen sowohl die EZ-Praxis als auch für die EZ-Evaluation vor zahlreiche Herausforderungen (ebd., Noltze & Harten, 2017).²³⁶ Kaplan, Vorwerk und Leiderer (2019, S. 18) sprechen gar von einer „Überforderung“: Durch die ambitionierten Ansprüche der Agenda werden Kernprinzipien der Wirksamkeitsagenda (u.a. Geberkoordination und -harmonisierung, *Ownership* der Partner*innen und *Alignment*) weniger umgesetzt und es kommt zu einer zunehmenden Bilateralisierung und (Re-)Fragmentierung in der Geberlandschaft, denn „bislang fehlt es den staatlichen EZ-Institutionen an Flexibilität und Anpassungsfähigkeit [...]“ (ebd., S. 19).²³⁷ Zwar finden Ziele und Prinzipien der Agenda 2030 zunehmend Einzug in politische Strategien, wie bspw. auch in Deutschland (vgl. BMZ, 2018), doch zahlreiche Länder und Bündnisse wie die EU hinken der Umsetzung hinterher (vgl. Kloke-Lesch & Messner, 2019, S. 14).²³⁸ Generell sind die existierenden M&E-Systeme zu schwach entwickelt, um den neuen Anforderungen genügen zu können (ebd.), was sich u.a. in den *freiwilligen Staatenberichten (Voluntary National Reviews, VNR)* zu den SDGs zeigt. Insgesamt gibt

²³⁵ Zur *Komplexität* s. Beiträge in Bamberger, Vaessen und Raimondo (2015). Zu Umsetzungsprinzipien s. Krempin (2019, S. 121): Die GIZ hat in Anlehnung an die Agenda 2030 fünf Prinzipien für die Umsetzung von EZ definiert: Gemeinsame Verantwortung; Integrierte Ansätze; Niemanden zurücklassen; Rechenschaftslegung und Universalität.

²³⁶ Zu Umbrüchen in der EZ-Praxis s. Klingebiel (2017) und die Beiträge in Sangmeister und Wagner (2019): z.B. Maihold (2019) bzgl. des Reformbedarfs der dt. EZ; Klingebiel (2019) zu Rahmenbedingungen (u.a. Rückkehr nationaler Interessen, Migrationsdruck); Krempin (2019) aus Sicht der GIZ.

²³⁷ Nach Klingebiel (2019, S. 48) gilt „die für viele Jahre prägende Debatte über aid effectiveness, [...] zwar weiterhin, verfügt aber kaum noch an über ein Momentum und internationale Fürsprecher“.

²³⁸ Die EU möchte laut eigenen Angaben eine aktive und vorantreibende Rolle hinsichtlich der SDG-Fortschritte übernehmen (vgl. European Commission, 2016, 739). Um hierzu regelmäßig zu berichten, wurde ein vom Umfang reduziertes *EU SDG Indikatoren Set* entwickelt, zu dem jährlich berichtet wird (vgl. Eurostat, 2018; 2019).

es zu wenige Daten; zu zwei Dritteln der Indikatoren kann derzeit nicht berichtet werden, weswegen eine sog. *Datenrevolution* gefordert wird (Rudolph & Holzapfel, 2017; Schwegmann & Holzapfel, 2018).²³⁹

Nachdem UNEG sich 2013 dem Netzwerk EvalPartners angeschlossen hatte, wurde 2015 zum *Jahr der Evaluation* erklärt (UNEG, 2013, 14. Mai). Auch wenn es laut Stockmann (CEval, 2015, S. 2) „(fast) keiner [gemerkt] hat“, wurden national und international zahlreiche Konferenzen und weitere Aktivitäten durchgeführt.

Standardisierung

Mit der internationalen Orientierung an der Agenda 2030 wurde auch eine inhaltliche Ausrichtung bestehender Standards unumgänglich: Als Ergebnis einer Diskussion um die Angemessenheit und Nützlichkeit der bestehenden fünf DAC-Kriterien wurden diese Ende 2019 um das Kriterium der *Kohärenz* ergänzt (DEval, 2019, 17. Dezember) – mit dem Ziel „[...] to better capture linkages, systems thinking, partnership dynamics, and complexity“ (OECD, 2019c, S. 3). Dass hierbei auch Akteur*innen der dt. EZ mitwirkten, zeigt sich daran, dass das Kriterium in den bereits seit 2006 bestehenden BMZ-Leitlinien (vgl. BMZ, 2006) vertreten ist. Das DEval erarbeitete mit dem BMZ, seinen Durchführungsorganisationen und dem DIE angepasste Leitlinien (vgl. BMZ, DEval & DIE, September 2020), d.h. die Prüffragen zu den Evaluationskriterien wurden geschärft und priorisiert (Faust & Verspohl, 2018; Noltze, 2019; Noltze & Hoffmann, 2019). Mit der Überarbeitung sollte auch die verbindliche Nutzung der deutschen Standards erhöht werden (ebd.).

Die Überforderung der staatlichen und nicht-staatlichen Organisationen, die sich an den SDGs orientieren, haben sich unterschiedliche Akteur*innen zu Nutze gemacht: Zahlreiche Handbücher, Leitfäden und Online-Tools geben Empfehlungen, wie man die SDGs evaluieren kann.²⁴⁰ Auf Empfehlung der OECD (2016, S. 131) erarbeitete das DEval in

²³⁹ Dies beinhaltet die Schließung von Datenlücken durch neue Technologien, die globale Stärkung der „Data literacy“ und die Entwicklung eines sog. *Datenökosystems*, das globale Standards schafft, die Datenqualität verbessert, Datenaggregationen ermöglicht und Datenmissbrauch verhindert (Schwegmann & Holzapfel, 2018). Zum Stand der Datenrevolution in Partnerländern s. ebd.

²⁴⁰ Bspw. Bamberger, Segone und Tateossian (2016) i.A. der *EvalPartners*; Als Online-Tools sog. *SDG Impact Assessment Tools* (Gothenburg Centre for Sustainable Development & University of Gothenburg,

Anlehnung an bestehende Standards und Richtlinien im Jahr 2018 neue *Evaluationsstandards* (vgl. DEval, 2018), in denen u.a. der DeGEval-Standard zu *Fairness um Unabhängigkeit und Integrität* erweitert wird (ebd., S. 2).²⁴¹ Die GIZ (GIZ Stabsstelle Evaluierung, 2018a, S. 11) gibt in einer aktuelleren Veröffentlichung an, sich an den Standards der DeGEval (2016) und an den Qualitätsstandards für die Entwicklungsevaluierung der OECD/DAC (OECD, 2010b) zu orientieren.

Wegweisende Studien

International wurde 2015 durch die UNEG eine wegweisende Studie durchgeführt, die eine *Review* der Evaluationserfahrungen der MDGs, eine Untersuchung zur Evaluierbarkeit der Agenda 2030 und daraus abgeleitete Empfehlungen beinhaltet (vgl. Aiazzi, 2016a, 2016b). Die OECD veröffentlicht alle zwei Jahre umfangreiche Studien unter dem Titel „Measuring Distance to the SDG Targets – An Assessment of where OECD Countries Stand“, in der die Distanz ausgewählter OECD-Länder zur Erreichung der SDGs dargestellt wird (vgl. OECD, 2019b).²⁴² Die OECD möchte hiermit ihre Mitgliedsländer dabei unterstützen, Stärken und Schwächen auf ihrem Weg zur Erreichung der Ziele zu identifizieren, womit sie sich bspw. von den *freiwilligen Staatenberichten (VNR)* unterscheiden (OECD, 2019b). Diese *nationalen Reviews* unterscheiden sich hinsichtlich ihrer Art, Form, Umfang und Qualität deutlich voneinander. Zudem mangelt es an einer klaren Reflexion und eines Verständnisses darüber, wie Evaluation in die nationalen Follow-up- und Reviewprozesse eingebunden werden soll (Meyer et al., 2018; Martinuzzi & Meyer, 2017).

Eine der ersten großen Studien der DEval sollte in Anlehnung an die Vorgängerstudien (Borrmann et al., 1999, 2001; Borrmann & Stockmann 2009a, 2009b) die Evaluierungspraxis der dt. EZ erneut untersuchen (vgl. Lücking, Freud & Bettighofer, 2015b). Die Ergebnisse zeigen, dass sich die Systemprüfungen in dem Sinne bewährt haben, da sie das

o.D.), auch angeboten von Beratungsfirmen wie Ramboll-Management Consulting (2018) oder S&P Trucost Limited (2021).

²⁴¹ Die Standards der DEval geben die Standards der DeGEval (2016) und die Qualitätsstandards der OECD (2010b) als Quellen an (vgl. DEval, 2018, S. 4).

²⁴² Aufgrund der Datengrundlage kann für Deutschland zu 103 der 169 SDG-Ziele eine Aussage getroffen werden. 20 der Ziele sind bereits erreicht. Deutschland schneidet bei den Zielen bzgl. des Zugangs zu Elektrizität, der Arbeitslosenquote und des Recyclings gut ab, ist allerdings bzgl. der Adipositasrate, des Tabakkonsums und der rechtlichen Rahmung für die Gleichstellung der Geschlechter noch weit von den Zielerfüllungen entfernt (OECD, 2019b).

deutsche EZ-System gestärkt haben: Neben dem Ausbau und der Weiterentwicklung in den EZ-Organisationen wird vor allem die Gründung des DEval als Erfolg bewertet. Nach wie vor besteht *Handlungsbedarf* hinsichtlich der Durchführung und „Nutzung von Wirkungs-, projektübergreifenden und ex-post-Evaluierungen“, wobei vor allem Evaluationsdesigns und genutzte Methoden gestärkt werden müssen (Lücking, Bettighofer & Freund, 2016, S. 4). Resümierend propagiert das DEval „eine politisch gewollte Kultur des ‚transparenten Lernens‘ [...], in der auch kritische Evaluierungsergebnisse stärker als bisher veröffentlicht und diskutiert werden“ (ebd., S. 4).²⁴³

Dass die *Evaluation nachhaltiger Entwicklung* zunehmend an Bedeutung gewinnt und künftig gewinnen wird (vgl. Martinuzzi & Meyer, 2017, S. 93), zeigt sich in zwei weiteren Evaluationen des DEval: Ergebnisse einer *Meta-Evaluation* (vgl. Noltze et al., 2018a), i.R. derer 513 Evaluationsberichte der KfW und GIZ als Stichprobe ausgewertet wurden, zeigen, dass „die Bewertung von Nachhaltigkeit [...] in der Praxis sehr uneinheitlich und unsystematisch [erfolgt]“ (Euler et al., 2018, S. 1). Basierend darauf empfiehlt die DEval eine „konzeptionelle Neuausrichtung der Evaluierung von Nachhaltigkeit“, die in allen EZ-Organisationen einheitlich und systematisch umgesetzt werden soll (ebd.). Basierend auf derselben Datengrundlage wurde eine weitere *Evaluationssynthese* veröffentlicht, die die *Einflussfaktoren auf die Nachhaltigkeitsbewertung* analysiert (vgl. Noltze et al., 2018b). Hier wird ersichtlich, dass die Kriterien zur Bewertung sehr unterschiedlich sind und dass die Art der Studie einen Einfluss auf die Bewertung von Nachhaltigkeit hat. Bspw. bewerten ex-post Evaluationen die Nachhaltigkeit von Vorhaben tendenziell am schlechtesten (ebd.). Die bis dato geltende konzeptionelle Grundlage, der BMZ-Leitfaden von 2006 (vgl. BMZ, 2006), greift in der Abbildung des Nachhaltigkeitsbegriffs zu kurz und ist daher überholt (Euler et al., 2018).

Eine weitere DEval-Studie zur Agenda 2030 aus Sicht der Bevölkerung (vgl. Schneider & Gleser, 2018) zeigte, dass die EZ zwar großen Zuspruch findet, allerdings große *Zweifel an der Wirksamkeit* bestehen. Zudem gibt über die Hälfte der Befragten an, von den

²⁴³Zudem wurde eine Baselinestudie zur Ausgangslage des DEval und seines Umfeldes durchgeführt (vgl. Lücking, Bettighofer & Freund, 2015a), mit dem Ziel können „[...] Veränderungen und Wirkungen, die auf das Institut zurückzuführen sind, über die Zeit“ zu messen und zu beschreiben (ebd., S. vii). Weitere Studien sind auf der DEval-Homepage gelistet.

SDGs bislang weder etwas gehört oder gelesen zu haben (vgl. Schneider & Eger, 2019, S. 64f), was vor allem auf Defizite i.R. entwicklungspolitischer Bildungsarbeit hinweist.²⁴⁴ Trotz der durchgeführten Studien lag Deutschland bei der letzten OECD Review bzgl. ihrer Ausgaben für unabhängige Evaluationen unter dem Durchschnitt (vgl. OECD, 2016, S. 133).

Schwerpunkte und Ansätze

Die Kombination und Integration verschiedener Methoden (*Mixed Methods*) ist nach wie vor noch aktuell und vor allem in der internationalen Evaluationsgemeinschaft weitgehend etabliert (vgl. Baur et al., 2017, S. 31).²⁴⁵ Bei der Weiterentwicklung von Standards und Methoden wird die dt. EZ seitens des DEval unterstützt (DEval, 2019). Sie bereitet auch aktuelle Diskussionen rund um Ansätze, Schwerpunkte und Methoden i.R. verschiedener Formate auf (bspw. in Form von *Policy Briefs*). Mit den sog. *Länderportfolioreviews* wurde bspw. ein Instrument geschaffen, wodurch „[...] Länderportfolios vor dem Hintergrund der Agenda 2030 unabhängig bewertet und steuerungsrelevante Informationen hinsichtlich der übergeordneten Relevanz, Kohärenz und Komplementarität der deutschen bilateralen EZ bereitstellt“ werden können (Verspohl & Hartmann, 2019, S. 23).

Die aktuelle *Debatte zur Wirkungsorientierung* in der EZ reicht in deutschen Evaluationskreisen von der Position, sie habe einen „nicht intendierten, hemmenden Einfluss“ (vgl. Koy, 2018) hin zum Ruf nach „mehr Wirkungsorientierung“ (vgl. Harten & Petrova, 2018). Laut Roxin (2016, S. 17) wird sich „die alte Frage nach der Wirksamkeit von EZ [...] weiter stellen, jedoch sind die Instrumente zur Beantwortung der Frage mit den jüngsten Veränderungen im Evaluierungsbereich geschärft“. Bspw. haben Ansätze, die *Kausalmechanismen* erklären, wie die *Kontributionsanalyse* oder *Process-Tracing*, in den vergangenen Jahren vermehrt Einzug in die Evaluationen der DEval gefunden (Schmitt,

²⁴⁴ Roxin (2016, S. 17) ist der Ansicht, dass Evaluation eine wichtige Grundlage in der entwicklungspolitischen Bildungsarbeit ist, um „[...] die Komplexität des Politikfeldes zu transportieren und mit einem ebenso geschärften Blick, die Wirkung entwicklungspolitischer Projekte kritisch zu hinterfragen“.

²⁴⁵ Anders als in Deutschland existieren international bereits zahlreiche Lehr- und Handbücher, eine eigene Fachgesellschaft und eine Fachzeitschrift (Baur et al., 2017).

2018).²⁴⁶ Der Nachfrage nach *rigorosen Wirkungsevaluationen* wurde u.a. durch die Verleihung des Nobelpreises für Wirtschaftswissenschaften 2019 an Esther Duflo, Abhijit Banerjee und Michael Kremer verstärkt (Bruder et al., 2019).²⁴⁷ Die GIZ hat bisher in ca. 40 Maßnahmen (quasi-)experimentelle Verfahren verwendet (vgl. Florian, Krisch, Till & Herrmanns, 2019), die KfW hat rigorose Methoden „erprobt“, einzelne NGOs haben erste Evaluationen durchgeführt (Bruder et al., 2019). Expert*innen kritisieren die in Europa und den USA zunehmenden rigorosen Analysen, bei denen *Big Data* (Massendaten) mit komplexen Algorithmen berechnet werden, die für Entscheidungsträger oder die Zivilgesellschaft schwer nachvollziehbar sind, weswegen öffentliche Grundsatzdiskussionen zu den Verfahren häufig außen vor bleiben (Martinuzzi und Meyer, 2017). Fakt ist jedoch, dass immer mehr und neue Daten generiert (bspw. Geo-, Satelliten- oder *Social-Media*-Daten) und gesammelt werden (Müller, 2017), die auch eine wichtige Rolle bei der weiteren Entwicklung von Evaluation haben werden (vgl. Noltze & Harten, 2017, S. 16; DEval, 2019, S. 99). Dadurch können u.a. *mehr Evaluationssynthesen* stattfinden, wie bspw. die Meta-Synthese der UNESCO zu SDG 4 (Bildung), bei der ca. 150 Evaluationen verschiedener Organisationen einbezogen wurden (vgl. Mundy & Proulx, 2019). An der Gemeinschaftsevaluierung (*Joint Evaluation*) waren mehr als 40 UN-Organisationen, NGOs und bilaterale Geber beteiligt (Frueh, 2019).

Durch die Vorgabe der Agenda 2030 von über 230 Indikatoren arbeiten EZ-Organisationen mehr denn je mit *Standardindikatoren* i.R. von *RBM-Modellen*, die letztendlich in eine aggregierte Berichtserstattung münden (vgl. Karkoschka et al., 2017).²⁴⁸ Die GIZ nutzt zur „Kommunikation mit der Öffentlichkeit und [dem] Dialog mit den Auftraggebern“ mehrheitlich *quantitative Aggregationsindikatoren* auf unterschiedlichen Ebenen (*Output, Outcome, Impact*), um sog. *Wirkungsdaten* zu verschiedenen Sektoren und SDGs zu generieren (vgl. Karkoschka et al., 2017). Die rein quantitative Berichterstattung wurde um qualitative Ergebnisse (vgl. Karkoschka, 2017) ergänzt. Meyer (2020, S. 234)

²⁴⁶ Zu Kontributionsanalysen s. Noltze und Schwedersky (2017), zu Process Tracing Orth (2019).

²⁴⁷ Die Forschenden belegten mit ihrer Arbeit die Bedeutung rigoroser Methoden im Bereich der Armutsbekämpfung und leisteten einen wichtigen Beitrag zur Etablierung dieser Methoden in der Entwicklungsforschung (Bruder et al., 2019).

²⁴⁸ Die *Asian Development Bank (ADB)* arbeitet als erste Organisation seit 2008 mit Standardindikatoren zu bestimmten Themen. Andere Organisationen folgten. S. hierzu Karkoschka et al. (2017, S. 8).

kritisiert in diesem Zusammenhang, dass Staaten nach fünf Jahren kaum über ihr „silo-thinking“ hinausgehen und lediglich endlose Indikatorenlisten für ihr Monitoring und „program-oriented accountability evaluations“ nutzen.

Neben *Unternehmensstrategischen Evaluierungen*, „[die] je nach Entscheidungs- und Erkenntnisbedarf vor Beginn der Maßnahmen angesetzte ex-ante Evaluierungen, begleitende (formative) Evaluierungen, ex-post Evaluierungen oder Querschnittauswertungen (Evaluierungssynthesen, Meta-Evaluierungen) sein [können]“, und Unternehmensentwicklung sowie Steigerung der Effektivität zum Ziel haben (GIZ Stabsstelle Evaluierung, 2018c, S. 4), führte die GIZ mit einer Reformierung ihres Evaluierungssystems²⁴⁹ *Zentrale Projektevaluierungen* ein (ebd., 2018a). Sie werden ex-post (zwei bis fünf Jahre nach Projektende) in Form von Schlussevaluationen (ca. acht Monate nach Projektende) oder auch als Zwischenevaluierung durchgeführt (ca. 12 Monate vor Projektende) und erlauben einen „kritischen analytischen Rückblick auf die Wirkungen und die Implementierung eines Vorhabens“ (GIZ Stabsstelle Evaluierung, Dezember 2020, 2018b, S. 6). Allerdings fanden Projektevaluierungen, die 2017 und 2018 durchgeführt wurden, alle noch zum Finanzierungsende statt, *keine davon ex-post* (vgl. Krämer, Almqvist, Raetzell, Alber & Verfürden, 2019). Im Zuge der Reformen veröffentlichte die GIZ auch eine umfangreiche „Theorie des Wandels“. Das systemische Wirkungsmodell beinhaltet Evaluationsaktivitäten sowie die Erläuterung von Wirkungsbereichen auf Outcome- und Impact-Ebene (vgl. ebd., 2018d).

Institutionalisierung und Professionalisierung

Die angesprochene *Reformierung des GIZ-Evaluierungssystems* Ende 2016 hatte zum Ziel, Wirkungsnachweise zu verbessern, die Glaubwürdigkeit von Ergebnissen zu erhöhen und Projektevaluierungen auf neue Herausforderungen (Agenda 2030) auszurichten. Als konkrete Maßnahmen zählen eine *neue Evaluierungs-Policy* (vgl. GIZ Stabsstelle Evaluierung, 2018e), eine verbreitetere Veröffentlichungspraxis von Evaluationsberich-

²⁴⁹ S. hierzu Ausführungen unter *Institutionalisierung und Professionalisierung*.

ten sowie die *Trennung von „Prüfung und Evaluierung“*, d.h. es gibt fortan eine Stabsstelle für Monitoring und eine weitere für Evaluation.²⁵⁰ Ein weiterer Weg der dt. EZ, den komplexen Herausforderungen der Agenda 2030 zu begegnen, sind sog. *Multiakteurspartnerschaften (MAP)*, bei denen mit staatlichen und wirtschaftlichen Akteur*innen sowie Akteur*innen aus der Wissenschaft und Zivilgesellschaft gemeinsame Lösungen gefunden werden sollen (Treichel, Höh, Biermann & Conze, 2016; GIZ, 2017a).²⁵¹ Die Online-Plattform *Partnerschaften 2030*, bei der die GIZ die Funktion eines Sekretariats innehat, wurde eingerichtet, um erfolgreiche Multiakteurspartnerschaften zu fördern und sie bei der Umsetzung ihrer Ziele zu unterstützen (s. *Partnerschaften 2030*, o.D.).

Nach gewissen Startschwierigkeiten, von Hauser (2017, S. 9) auch als „Wachstumsstörungen“ bezeichnet, erfolgte ab Mitte 2014 ein ‚*Neustart‘ des DEval*. Ab dann konnten sich die Mitarbeiter*innen u.a. Themen wie Methodenforschung und Wirkungsanalysen sowie der Bedeutung der Stärkung von Evaluationskapazitäten widmen (ebd.). Nach Stockmann und Meyer (2017a) hat die Gründung des DEval allerdings vorerst nicht zu mehr, sondern zu weniger Evaluationen in der EZ geführt, da mit der Institutsgründung BMZ-Evaluationen wegfielen und die GIZ ihre unabhängigen Evaluationen weitgehend einstellte (ebd.).²⁵² Die OECD (2018, 14. Dezember, S. 4) sieht in der Institutsgründung einen großen Mehrwert und empfiehlt, die Erfahrungen in einem größeren Umfang wie bislang zu teilen. Neben der Durchführung von Studien zum Thema Nachhaltigkeit²⁵³ beteiligt sich das Institut 2019 an der Ausrichtung der 22. Jahrestagung der DeGEval zum Thema *Evaluation und Nachhaltigkeit*, bei der beleuchtet wurde, „[...] wie Evaluation mit

²⁵⁰ Auch wenn sich die Nutzung der Begriffe im Laufe der Zeit verändert hat, bezeichnet die GIZ (o.D.a) aktuell „das ‚Wirkungsorientierte Monitoring‘ [als] eine wesentliche Voraussetzung für Evaluierungen während des Projekts und nach Projektende“. In ihrer *Policy für Monitoring und Evaluierung* von 2013, die unter der damaligen *Stabsstelle für Monitoring und Evaluierung* veröffentlicht wurde (vgl. Schweitzer, 2013), betont die GIZ die Verknüpfung zwischen den beiden Komponenten (M&E). Auf die „optimale Verknüpfung von Monitoring und Evaluierung [...] als Merkmal des Auftragsmanagements der GIZ“ wird in der aktuellen Evaluierungspolicy zwar hingewiesen, allerdings nicht weiter ausgeführt (vgl. GIZ Stabsstelle Evaluierung, 2018e, S. 14).

²⁵¹ Treichel et al. (2016) geben einen Überblick zu Definitionen und Typisierungen von MAPs.

²⁵² Nach fünf Jahren verzeichnete die DEval 18 Berichte (DEval, 2017), in den Jahren 2017-2019 wurden 15 Berichte und 20 Policy Briefs veröffentlicht (DEval, 2019).

²⁵³ S. Ausführungen unter *Wegweisende Studien*.

Nachhaltigkeit umgehen kann, wie sie gemessen werden kann, welche Herausforderungen dabei bestehen und welchen Beitrag unsere Profession zu einer nachhaltigen gesellschaftlichen Entwicklung leisten kann“ (Flatters & Pütz, 2020, S. 186).

Der *Aufbau von Evaluationskapazitäten*²⁵⁴ als Teil des Professionalisierungsprozesses ist weiterhin von großer Relevanz, was bspw. die Umsetzung umfangreicher Projekte in Uganda und in Lateinamerika demonstrieren, an denen auch die DEval beteiligt ist (vgl. Geuder-Jilg, Stadtmüller & Harten, 2019; Stockmann, 2018, S. 367). Lehr- und Lernformate wie das *Blended Learning* oder *E-Learning* spielen eine immer größere Rolle; der Markt der Weiterbildungsangebote zur Evaluation, die online zur Verfügung stehen, wächst (vgl. Nolte & Schopper, 2017, S. 139). EZ-bezogene Netzwerke und Organisationen wie bspw. *EvalPartners* oder die UN-Welternährungsorganisation bieten im Bereich M&E Online-Kurse an (vgl. ebd., S. 136f). Auch Beratungsfirmen wie bspw. *inFocus* entdecken diesen Markt für sich (vgl. *inFocus*, o.D.). Generell spezialisieren sich zunehmend private Firmen im Bereich M&E und agieren als externe Gutachtende bzw. Evaluierende z.B. für die GIZ (u.a. PHINEO, GOPA Worldwide Consultants, GFA Consulting Group, Mainlevel Consulting AG).²⁵⁵ Allerdings besteht auch ein Mangel an Evaluationen, bspw. hinsichtlich der Wirksamkeit einzelner Formate, sowie unzureichende Definitionen und Konzepte (Stockmann, 2018).²⁵⁶

Im Hinblick auf die die *Veröffentlichungspraxis* deutscher EZ-Akteur*innen hat sich die bislang konsequente „Abschottung gegenüber öffentlicher Kritik“ (Borrmann et al., 1999, S. 26) im Laufe der Zeit gewandelt: So berichtet bspw. die GIZ zu Projekterfolgen und -misserfolgen (vgl. GIZ, 2017, S. 22f.) und zu Schwächen in der Qualität ihrer Evaluationen (vgl. ebd., S. 26ff.). Das BMZ verweist umfangreich auf die Evaluationen der DEval und veröffentlicht eigene Stellungnahmen zu allen Evaluationsberichten des Instituts (vgl. BMZ, 2010-2020c).

²⁵⁴ Häufig auch unter dem Begriff *Evaluation Capacity Development (ECD)* zu finden.

²⁵⁵ Die Aufzählung erhebt keinen Anspruch auf Vollständigkeit. Das von PHINEO entwickelte *Kursbuch Wirkung* (vgl. Kurz & Kubek, 2017) wird auch in dieser Arbeit genutzt.

²⁵⁶ Basierend auf dem *CEVAL-Ansatz* zur Wirkungsevaluation von Projekten und Programmen (vgl. Stockmann, 2009) entwirft Stockmann (2018) einen *CEval-ECD-Ansatz*, der im Vergleich zu anderen Ansätzen systemisch angelegt ist (vgl. ebd., S. 388f).

Auf internationaler Ebene wurde bereits 2011 in Busan eine neue *Globale Partnerschaft für wirksamere EZ (Global Partnership for Effective Development Cooperation – GPEDC)* beschlossen (vgl. Busan HLF, 2011, Dezember). GPEDC vereinigt zahlreiche staatliche, nichtstaatliche und privatwirtschaftliche Akteur*innen (GPEDC, o.D.) und beschreibt ein Monitoring von zehn Indikatoren. Die Daten, die sich aus nationalen Datensammlungen speisen, tragen auch zur Berichterstattung zu einzelnen *SDGs* bei (OECD/UNDP, 2019, S. 14). Sie werden in regelmäßigen Abständen evaluiert und stehen über eine Homepage der Öffentlichkeit zur Verfügung.²⁵⁷ Das seit 2013 bestehende *High-level Political Forum on Sustainable Development (HLPF)* der UN übernimmt eine zentrale Rolle hinsichtlich des Follow-ups und der Überprüfung der *SDGs*. Die freiwilligen Staatenberichte dienen als Grundlage für die regelmäßigen Reviews des HLPF (vgl. UN DESA, 2019). Hierdurch entstand „[...] eine positive Dynamik von kollektivem Druck und gegenseitigem Lernen [...]“ (Kloke-Lesch & Messner, 2019, S. 14). Zudem wird in verschiedenen Sprachen ein Handbuch für Länder angeboten, die ihre nationalen Berichte vorbereiten (vgl. UN DESA, 2019, November).

Im Zuge der Einführung der *SDGs* etablierte *EvalPartners* das Netzwerk *EvalSDGs* „[...] in response to the need to embed effective monitoring and evaluation in the *SDGs*“ (EvalSDGs, o.D.). Auf der Homepage bietet die globale Arbeitsgruppe zahlreiche *Briefing Papers*, Publikationen, nationale Erfahrungsberichte und *Insights* rund um die *SDGs* (s. EvalSDGs, 2021). EvalSDGs ist ein Ergebnis des Jahrs der Evaluation 2015 und begleitet die in diesem Kontext beschlossene *Global Evaluation Agenda 2016-2020*.²⁵⁸ Die bereits 2013 gegründete Arbeitsgruppe *Evaluating Sustainable Development* der European Evaluation Society arbeitet eng mit EvalSDGs zusammen (Martinuzzi & Meyer, 2017). Aufgrund derartiger, seitens der UN vorangetriebener Initiativen besteht „[...] durchaus Hoffnung, dass die Evaluation nachhaltiger zumindest auf globaler Ebene in Verbindung mit den *SDGs* besser verankert werden kann“ (ebd., S. 89).

²⁵⁷ S. UNDP (2022). Ausführlich zum Hintergrund und Gründung der GPEDC s. Abdel-Malek (2015).

²⁵⁸ 2014 startete EvalPartners, u.a. in Kooperation mit IDEAS, IOCE und UNEG, einen globalen Beratungsprozess. Ergebnis war die Definition von vier Kernbereichen der *Global Evaluation Agenda 2016-2020* (auch *EvalAgenda2020* genannt) (1. Enabling environment for evaluation; 2. Institutional capacities for evaluation; 3. Individual capacities for evaluation 4. Inter-linkages among the first three dimensions) und dazu gehörige Aktionen (vgl. EvalPartners, 2016, 15. Januar).

4.6 Evaluation in der EZ – ein Zwischenfazit

Auch wenn Evaluation in der Entwicklungspolitik im Vergleich zu anderen Politikfeldern bereits gut verankert zu sein scheint (vgl. Wessels, 2020), „hinkt“ die Evaluationskultur in Deutschland im internationalen Vergleich „hinterher“, der Stand der Professionalisierung „lässt noch sehr zu wünschen übrig“ (Gomez, 2017, S. 159f.). Der Leiter der Stabsstelle für Evaluierung der GIZ sieht vor allem die *DeGEval* in einer zentralen Rolle bei der Professionalisierung und „bei der *Entwicklung eines Marktes für Evaluation in Deutschland*“²⁵⁹ (ebd., S. 161ff.). Stockmann und Meyer (2017a) identifizieren das politische System zwar als maßgeblichen Treiber der Evaluationsentwicklung in Deutschland, bemängeln jedoch das Fehlen einer bundeseinheitlichen Strategie oder einer gesetzlichen Grundlage, wie es bereits in anderen Ländern der Fall ist (vgl. ebd., S. 72). Zudem sprechen sie von einer „Stagnation“ der Evaluation im Wissenschaftssystem (ebd., S. 58).²⁶⁰

Zwar verzeichnen die *Systemprüfungen des deutschen EZ-Evaluationssystems* seit ihrer ersten Initiierung Ende der 1990er Jahre Fortschritte, es besteht jedoch nach wie vor Handlungsbedarf hinsichtlich der Durchführung und „Nutzung von Wirkungs-, projektübergreifenden und ex-post Evaluierungen“, wobei vor allem Evaluationsdesigns und genutzte Methoden gestärkt werden müssen (Lücking, Bettighofer & Freund, 2016, S. 4). Durch die Anstrengungen der OECD wurden zudem maßgebliche Weiterentwicklungen hinsichtlich der *Standardisierungsprozesse* erzielt, was auch *Definitionen zu Evaluationsterminologien* beinhaltet. Dennoch kann nicht ausgeschlossen werden, dass Begriffe immer noch uneinheitlich oder unzureichend genutzt werden. Vor allem der *Nachhaltigkeitsbegriff* wurde in den neuen BMZ-Orientierungslinien in einer Neuauflage von 2021 zwar um neue Prüffragen erweitert, er geht allerdings über die Dimension der Dauerhaftigkeit nicht hinaus (vgl. BMZ, 2021, S. 11). Neben den bisherigen Kriterien sind die Orientierungslinien um Prüffragen zu den *Beiträgen einer Maßnahme zur Agenda 2030*

²⁵⁹ Hervorhebungen durch die Autorin.

²⁶⁰ Stockmann und Meyer (2017a) evaluieren die Entwicklung der Institutionalisierung und Professionalisierung von Evaluation in Deutschland i.R. dreier Subsysteme (politisches, soziales und Wissenschaftssystem) und skizzieren für die Zukunft unterschiedliche Szenarien (s. ebd. 102)

für *nachhaltige Entwicklung* ergänzt worden (vgl. ebd., S. 12f.). Die *Evaluation nachhaltiger Entwicklung*²⁶¹ und der *SDGs* wurde seitens der „globalen Evaluationsgemeinschaft“ schnell konzeptionell bedacht, ist in den Länderstrategien zur Umsetzung der *SDGs* jedoch mit einer zeitlichen Verzögerung angekommen (vgl. BMZ, 2018; Martinuzzi & Meyer, 2017, S. 88f.).

Mit der Einführung der Agenda 2030 verzeichnet Stockmann (2017) eine *Zunahme an Funktionen und Aufgaben von Evaluation*, v.a. bzgl. ihrer *Aufklärungsfunktion* (als Teil der *Erkenntnisfunktion*) im Hinblick auf die *SDGs*, die bislang neben der weit verbreiteten *Steuerungs-, Kontroll-* und der *Legitimitätsfunktion* am schwächsten entwickelt ist. Auch Feinstein (2019) spricht sich für eine Erweiterung der Rolle und Funktion von Evaluation aus und plädiert für die Nutzung eines *Konzepts der dynamischen Evaluation*, die u.a. die Nutzung verschiedener Methoden und Techniken, Aspekte des „Scaling-ups“ und „quasi real-time evaluation“ (durch die Nutzung von *big data*) beinhaltet: „The usual ways of formulating the roles of evaluation in terms of learning and accountability are no longer adequate. Evaluation and evaluators need to go beyond this traditional formulation and discussion if they are to effectively contribute to transformational change (ebd., S. 18).“

Die ‚alte Tradition‘ der Nutzung *theoriegeleiteter Logical Framework-Ansätze* zur Messung der Zielerreichung oder sog. Erfolgsfaktoren von Maßnahmen ist in Form des RBMs oder ähnlichen Modellen wie dem WoM der GIZ auch heute noch aktuell. Dem umfangreichen Zielsystem und der Zunahme an zur Verfügung stehenden Daten begegnen EZ-Organisationen u.a. mit dem Einsatz von meist quantitativen *Standardindikatoren*, die Datenaggregationen zulassen. Holzapfel (2014, S. 1) warnt in diesem Zusammenhang vor „[...] einer zu starken Fokussierung auf kurzfristige und [...] leicht abbildbare Ergebnisse, und [einer damit verbundenen] Vernachlässigung längerfristiger und qualitativer Wirkungen“. Eine aggregierte Berichterstattung reiche alleinig nicht aus, um die Wirk-

²⁶¹ Nach Stockmann und Meyer (2017a, S. 95) verschmelzen die Bereiche Politik und Umwelt im Konzept der nachhaltigen Entwicklung: „Mit Blick auf Evaluation bedeutet dies, dass die [...] langjährige Evaluationstradition der Entwicklungspolitik auf die ebenso entwickelte technische Monitoringpraxis der Umweltpolitik trifft und sich eventuell gegenseitig befruchten kann. Ob und wie dies erfolgen könnte ist zurzeit noch nicht absehbar.“

samkeit der Beiträge einer EZ-Organisation zu bewerten. Daher werden ergänzende Methoden (Einbeziehung qualitativer Ergebnisse, ergänzende rigorose Wirkungsevaluierungen) sowie die Verfolgung des Prinzips der *contribution* (Beitrag) statt der *attribution* (Zuschreibung) empfohlen (ebd.). Auch Martinuzzi und Meyer (2017, S. 93) warnen vor einer „methodischen einseitigen Fokussierung“. Sie sehen die *Qualitätskontrolle* sowie die *methodische Transparenz* bzgl. der Vor- und Nachteile eingesetzter Methoden (hier: rigorose Analysen mit sog. *Big Data*) als eine besondere Herausforderung für die künftige Evaluation (ebd, S. 93). Nach Hauck (2019) sollte der Fokus darauf liegen, die zahlreichen, sich stetig vermehrenden Informationsquellen ‚nutzbar zu machen‘, da die Auswertung von Sekundärquellen „schneller und kostengünstiger“ ist (vgl. ebd., S. 136).²⁶²

Aufgrund einer steigenden Nachfrage und der Evaluationspflichten bei öffentlich geförderten Maßnahmen wird eine *Zunahme durchgeführter Wirkungsevaluierungen* prognostiziert (Müller, 2017). Der Anspruch an Designs und die methodische Qualität von Wirkungsevaluierungen wird sich gemäß dem wissenschaftlichen Fortschritt weiter erhöhen. Dies zeichnet sich bisweilen in der steigenden *Nachfrage anspruchsvoller (experimenteller) Designs* wie dem ‚Goldstandard‘ der RCT auf Seiten der Auftraggebenden ab, was interne und externe Evaluierende in ihrer Methodenkompetenz fordert (ebd.). Da solche *rigorosen Wirkungsevaluierungen* in der dt. EZ bislang eher weniger durchgeführt wurden, empfiehlt die DEval ihre Verankerung im EZ-System, die Bereitstellung dafür benötigter Ressourcen, den Einbezug der Wissenschaft sowie den Ausbau von Evaluationskapazitäten sowohl in den EZ-Organisationen als auch in Partnerländern (Bruder et al., 2019).

Aufgrund der Zunahme an Daten wird auch ein *Anstieg an Evaluationssynthesen* (inkl. qualitativer Meta-Synthesen und quantitativer Meta-Analysen) prognostiziert, bei denen bspw. Ergebnisse mehrerer Evaluationen zu einem speziellen SDG zusammengetragen und untersucht werden. Allerdings bleibt die Umsetzung solcher sehr ressourcenintensiven Gemeinschaftsanalysen, wie die Meta-Analyse der UNESCO zu SDG 4, fraglich

²⁶² Sie hierzu ausführlich Müller (2017, S. 216) und Leeuw (2017). Letzterer argumentiert für die Nutzung von Big Data, dass sie bereits vorhanden ist und nicht aufwendig erhoben werden muss.

(vgl. Frueh, 2019, S. 29). Denn trotz der nicht neuen Forderung nach mehr Politikkohärenz und Geberharmonisierung zeigt sich neben der Abkehr von gemeinsamen programm-basierten Ansätzen auch eine rückläufige Anzahl von *Gemeinschaftsevaluationen* (*Joint Evaluations*) (DEval, 2019).²⁶³

Nachdem *partizipative Evaluationsansätze* nach anfänglichen Startschwierigkeiten schließlich in den 1990ern auf eine breitere Zustimmung trafen, und im weiteren Verlauf, v.a. in den 2000ern immer mehr Ansätze entwickelt wurden, erleben sie jüngst einen „Boom“ in Nordamerika, aber auch zunehmend in Europa (Meyer et al., 2017, S. 199). Das *Spannungsfeld zwischen Partizipation und Bewertung* hat weiter Bestand, ebenso wie damit verbundene methodische Herausforderungen. Bislang fehlen Studien, die verschiedene Ansätze vergleichen, oder Verfahren zur Bewertung der Partizipationsqualität (ebd.). Im Zusammenhang mit Partizipation steht das Prinzip „leave no one behind“ der Agenda 2030. Eine *inklusive Denkweise* i.R. von Evaluation, bei der gewährleistet wird, dass „keiner zurückgelassen wird“, nimmt demnach noch an Bedeutung zu (vgl. Meyer, 2020, S. 231).

Meyer (2020, S. 233) spricht von einer „großen Lücke“ zwischen den Fortschritten und Entwicklungen in der Evaluationsforschung, die sich der mit der Agenda 2030 verbundenen Herausforderungen annehmen und nach seiner Aussage auch gewachsen sind, und der Evaluationspraxis:

„On one hand side, more and more scientists show their ability to overcome traditional barriers in the academic system and provide interdisciplinary, inclusive, multi-dimensional, dynamic, mixed-method, practical but scientifically sound evaluation approaches. On the other hand, political practice (especially on the national level) seems to tend in the other direction. Nationalistic egoism, populism and simplified solutions are on the rise, neglecting global threats as well as the importance of international cooperation and the need to compromise for joint action“.

Nationale (bspw. Meyer, 2020; Vaessen, 2019) und internationale Evaluierende (bspw. Magro & van den Berg, 2019) sprechen sich für die *Nutzung systemischer Evaluations-*

²⁶³ Die *OECD Review* von 2016 merkt kritisch an, dass weniger als die Hälfte der untersuchten EZ-Organisationen die Rolle von Partnerländern, *Joint Evaluations* oder kollaborative Geber-Nehmer-Evaluationen in ihren Evaluationpolicies erwähnen (vgl. OECD, 2016, S. 10).

ansätze aus (Vaessen, 2019). Des Weiteren sollten *Ansätze aus der Nachhaltigkeitsforschung*, die sich mit der *Transformation*, d.h. der Umstellung von einem soziotechnischen System zu einem anderen befassen, genutzt werden. D.h. dass Evaluationskonzepte bspw. die Verzahnung zwischen globaler und lokaler Ebene bedenken, sektorenübergreifend angelegt sind (nicht auf einzelne SDGs fokussiert) und eine gewisse ‚adaptive Resilienz‘ aufweisen. Hauptproblem hierbei bleibt die Übersetzung einer langfristigen Denkweise (von Wissenschaftler*innen) in kurzfristige Lösungsansätze (von politischen Entscheidungsträgern) (Meyer, 2020, S. 228f.). Die Veröffentlichung *Blue Marble Evaluation* von Patton (2019) stellt in diesem Zusammenhang möglicherweise den Beginn einer neuen Generation der Evaluation dar: „Patton weist der Evaluation bzw. den Evaluierenden eine sehr weit gehende Funktion zu: zusammen mit den von ihnen evaluierten transnationalen Netzwerken eine globale Transformation mitgestalten“ (Beywl, 2020, S. 346). Hierzu formuliert Patton (2019) *Blue Marble Prinzipien* wie bspw. *Global-Denken* oder *Einbringen nach Transformationsbedarf* (Beywl, 2020, S. 344).

5. Betrachtung des SfE-Kontexts: Evaluation und Forschung

„The requirement and need for M&E has become one of the most significant challenges, and sites of critique, within the SDP sector in recent years.” (Collison et al., 2019, S. 6)

Bereits Definitionen rund um die Forschung und Evaluation sowie das Verständnis bzgl. der Evaluationsfunktionen, womit das folgende Kapitel startet, stellen im Kontext SfE eine Herausforderung dar und beinhalten kritische Debatten. In diesem Kapitel wird bewusst *Forschung* zusätzlich zur Evaluation benannt und dazwischen unterschieden, wo dies aus der Literatur ersichtlich oder als notwendig erachtet wird; so auch im Teilkapitel zu *Evaluierenden und Forschenden*, in dem es gilt, deren Rollenverständnis zu erörtern, da sich nicht alle Personen, die sich im SfE-Kontext wissenschaftlich betätigen, als Evaluierende bezeichnen. Im Anschluss wird die *Entwicklung des ‚Forschungsfeldes‘* nachgezeichnet, was alle Forschungs- und Evaluationsaktivitäten im SfE-Kontext einbezieht und wobei auf prägende Kontexte und Meilensteine Bezug genommen wird, Institutionalisierungs- und Professionalisierungsprozesse sowie Standardisierungsprozesse veranschaulicht werden. Nachdem die inhaltlichen Schwerpunkte des Forschungsfeldes beschrieben sind, wird aufgrund seiner disziplinären Diversität den theoretischen Zugängen ein separates Kapitel gewidmet. Im Anschluss werden genutzte Ansätze, inklusive Designs und Methoden sowie die allgegenwärtige Debatte um Wirkungen betrachtet, die ob der *‚Sonderstellung‘ von Sport im EZ-Kontext* („low-cost, high-impact tool“, Kay & Dudfield, 2013, S. 5) eine besonders wichtige Rolle spielt. Das Kapitel schließt mit einer kurzen Übersicht zu den Forschungs- und Evaluationsaktivitäten zu Sport in der dt. EZ.

5.1 Definitionen und Funktionsverständnis von Evaluation

In den vornehmlich englischsprachigen wissenschaftlichen Beiträgen im SfE-Kontext werden die Begriffe *research*, *evaluation* und *M&E* bunt genutzt, eine eindeutige begriffliche Abgrenzung findet häufig nicht statt.²⁶⁴ *‚(Conducting) Research‘* wird häufig mit

²⁶⁴ Ausnahmen, die eine klare Stellung zu begrifflichen Abgrenzungen beinhalten, sind bspw. Levermore (2011, S. 340), Welty Peachey und Cohen (2015, S. 17; 2016, S. 283) und Whitley et al. (2020, S. 22).

M&E-Aktivitäten gleichgesetzt, was zwangsläufig zu einem undifferenzierten (Meinungs-)Austausch führt (vgl. bspw. Kay, 2012). Welty Peachey und Cohen (2016, S. 283) verweisen auf die Notwendigkeit zur Unterscheidung:

„It must be noted that research and evaluation in SDP are not necessarily the same thing. Research can be undertaken to produce knowledge and understand a phenomenon (such as policy and program theory in SDP) without having a goal of evaluating or improving an SDP organization or program, whereas evaluation is undertaken with the goal to give something back to the organization to help it improve and may originate from a contract between researcher(s) and the organization.“

Sie merken auch an, dass sogar etablierte SfE-Wissenschaftler*innen *research* mit *programme evaluation* gleichsetzen und dies ein Problem für den SfE-Kontext darstellt: „This could be problematic for the SFD field if scholars assume that monitoring and evaluation is the most significant form of SFD research and neglect engaging in basic research and critical reflections for the sole purpose of advancing knowledge and building theory“ (Welty Peachey & Cohen, 2015, S. 24).

Für Jeanes und Lindsey (2014, S. 203, S. 199) kann M&E als „processes of research“ betrachtet werden. Sie sehen die Unterscheidung darin, dass M&E hauptsächlich im Programmkontext genutzt wird und auf die Optimierung abzielt, während *research* darüber hinausgeht (ebd., S. 201). Sie plädieren für eine genauere Unterscheidung der Begriffe *research*, *evaluation* und *monitoring*. Sie beziehen M&E hauptsächlich auf den Projekt- oder Programmkontext und denken, „a research-focused approach“ würde einen Beitrag dazu leisten „[...] to develop knowledge of wider social contexts and processes to develop [a] type of understanding [...] [which] would be valuable and also support improvements in practice“ (ebd., S. 213). Whitley et al. (2020) merken allerdings an, dass „[...] evaluation and research are not mutually exclusive, and there are arguments that evaluation is a subset of research (and vice versa)“ (ebd., S. 22). Auch Collison et al. (2019, S. 6) sind der Ansicht, dass sich *research* nicht von M&E-Prozessen trennen lässt:

„An additional point of contestation is whether conducting research and managing M&E are one and the same. We would argue that the assumption that research *necessarily* informs, guides, influences or even constructs M&E frameworks or evidence is misguided. Progressive research methodologies focused on M&E, for example participatory action research, may well serve the dual purpose of knowledge production while producing and assisting with the formal process of evidencing and reporting, but the relationship between these processes requires ongoing negotiation and reflexivity.“

Generell wird häufig die Nutzung der im Monitoringprozess gesammelten Daten zur Evaluation betont (vgl. Coalter, 2006, S. 9; Elsemann et al., o.J., S. 7). Monitoringaktivitäten werden meist als fester Bestandteil eines Programmzyklus verstanden und als eine Aktivität, die intern durch Programmmitarbeitende durchgeführt wird – mit dem Ziel, „[...] to keep track of what is happening and check progress being made towards achieving objectives“ (Levermore, 2011, S. 340). Wie im EZ-Kontext verschmelzen die beiden Begriffe auch im Bereich von SfE häufig zu einer Einheit, dabei erfordern und bedienen sie unterschiedliche Aspekte: „As they pose and respond to different types of questions, it is evident that monitoring and [...] evaluation [...] require different tools, different skills, different strategies and ideally different personnel“ (Kaufman, Rosenbauer & Moore, 2014, S. 177). Ein weiterer, neuer Trend im SfE-Kontext ist die Wortkreation *MEL* (*Monitoring, Evaluation, Learning*), was die internen Lernprozesse, die bspw. i.R. einer Organisation während der M&E-Prozesse stattfinden, hervorheben soll (vgl. Whitley et al., 2020, S. 22). Der Begriff betont demnach eine *Funktion* von Evaluation (*Lernfunktion*).

Die zweidimensionale Funktionszuschreibung von Evaluation in *accountability* (*Legitimierung*) und *learning* ist im sportbezogenen EZ-Kontext weit verbreitet, u.a., um auch die unterschiedlichen Absichten von Geldgebenden und Empfänger*innen zu verdeutlichen und deren Rollen zu polarisieren (vgl. Coalter, 2009, S. 70; Kay, 2012, S. 888; Levermore & Beacom, 2012). Kritisierende betonen die Vernachlässigung der *Lernfunktion*. Besonders M&E-Systeme, die hier klar von *research* abgegrenzt werden (s.o.), würden sich zu stark auf die externe Rechenschaftspflicht bzw. Legitimierung fokussieren und lokale Lernprozesse limitieren. Sie fordern Evaluationsprozesse, bei denen die *Lernfunktion* und das *Teilen von Wissen* (*Aufklärung, Dialog*) im Vordergrund steht (Coalter, 2009, S. 70; Kay, 2012, S. 888, 891). Coalter (2009, S. 71) spricht auch von der *Entwicklungsfunktion* im Zusammenhang mit M&E und bezieht sich hierbei hauptsächlich auf die Entwicklung von Organisationen.

5.2 Evaluierende und Forschende aus SfE-Perspektive

Die uneinheitliche Nutzung und die unterschiedliche Auffassung von Definitionen und Funktionen von Evaluation haben Auswirkungen auf das Rollenverständnis von Evaluierenden und Forschenden im SfE-Kontext.

Häufig wird im Zusammenhang mit Forschung und Evaluation im SfE-Kontext von „power dynamics that exist between Global North funding agencies, Global North researchers, Global South NGOs and on-the-ground-deliverers“ (Jeanes & Lindsey, 2014, S. 206) gesprochen und davon, dass *ungleiche Machtverhältnisse* „[...] particularly along lines of race, class and geography [...]“ Untersuchungsprozesse wie die Datensammlung und -interpretation beeinflussen können (Collison et al., 2016, S. 413). Vor allem werden die M&E-Praktiken der mehrheitlich *externen Wissenschaftler*innen aus dem Globalen Norden* kritisiert: Sie verbrachten mehr Zeit mit den Geberorganisationen, die die Forschung finanzieren, als mit den Organisationen vor Ort; d.h. zu wenige Tage vor Ort, um bspw. Zugang zu lokalen Projektkontexten zu erlangen, diese zu verstehen und Beziehungen zu lokalen Projektbeteiligten aufbauen zu können (Kay, 2009; 2012; Darnell, Chawansky, Marchesseault, Holmes & Hayhurst, 2018a; Nicholls, Giles & Sethna, 2011). Kay (2009, S. 1189) betont, dass die Anwendung von Evaluationsmethoden aus dem Globalen Norden, vor allem in Form von M&E-Systemen (*Logischen Modellen*), die Gefahr birgt, dass authentische, lokale Stimmen überhört oder einfach vergessen werden. Sie plädiert für „in-depth, reflexive approaches“ (ebd., 2011, S. 10) und partizipative Methoden, die lokale Kenntnisse der örtlichen Gegebenheiten einbeziehen (ebd., 2012, S. 899). Der Einsatz *qualitativer Forschungsansätze* positioniert den „researcher“ laut Darnell, Whitley und Massey (2016, S. 573) „[...] less of a detached evaluator and more of a collaborator“. Sugden, Adair, Schulenkorf und Frawley (2019, S. 277) reflektieren anhand eines Fallbeispiels die Herausforderungen, mit denen „outsider, researchers from the Global North“ i.R. von kurzfristigen ethnografischen Forschungsaufenthalten (*short-term ethnography – STE*) konfrontiert sind. Anders als Vertreter*innen von ausschließlich qualitativen Forschungsansätzen wie Kay sehen sie im Ansatz von STE auch Chancen. Sie empfehlen anderen Forschenden aus dem Globalen Norden, denen nur ein derartiger Forschungsaufenthalt möglich ist, eine vorherige ausführliche ethnographische Er-

kundung des Forschungskontexts, eine intensive Vorbereitung („reading deeply“, „seeking advice from academic colleagues“, „engaging with [local] mentors“), die Nutzung digitaler Kommunikation vor und nach dem Forschungsaufenthalt sowie die ständige Reflexion von und die eigene *Einbettung in postkoloniale Paradigmen* (ebd., S. 285).

Tatsache ist, dass die meisten Studien bis auf wenige Ausnahmen *extern* von Wissenschaftler*innen aus dem Globalen Norden in Form von *Fremdevaluierungen* durchgeführt und veröffentlicht werden und dass es SfE-Organisationen meist an finanziellen Ressourcen und Expertise mangelt, um auch *interne* Evaluationen durchführen zu können (Cronin, 2011, S. 33ff.; Schulenkorf et al., 2016, S. 29f.; van Eekeren, 2013, S. 35; Welty Peachey & Cohen, 2016). Auch wenn *research collaborations/partnerships* sowohl von Seite der Forschenden bzw. Evaluierenden (Darnell, Giulianotti, Howe & Collison, 2018b; Kaufman et al., 2014; Welty Peachey & Cohen, 2016; Whitley, Farrell, Wolff & Hillyer, 2019b; Whitley et al., 2020) als auch von Seiten der SfE-Praktizierenden (Shin, Cohen & Welty Peachey, 2020) als wichtig und gewinnbringend erachtet werden und sie die gemeinsame Wissensgenerierung („co-creation of knowledge“) durch „researcher, practitioner, and participant“ als logischen und unabdingbaren Schritt im Bereich von SfE ansehen (Darnell et al., 2018b, S. 145), verweisen sie zeitgleich auf umfangreiche *Herausforderungen und Barrieren*, dies umzusetzen (Welty Peachey & Cohen, 2015, 2016).²⁶⁵ Sie gestehen ein, dass „Forschende aus dem globalen Norden oft mehr aus der Forschungsbegegnung und -beziehung mitnehmen, als dass sie geben“ (Collison et al., 2016, S. 419) und sprechen von der „*Isolierung der Wissenschaft*“ (Whitley et al., 2019b, S. 8).²⁶⁶ Strategische Partnerschaften, wovon alle Beteiligten profitieren würden, die Kultivierung eines gemeinsamen Verständnisses (z.B. Zielsetzungen, Einbindung in alle Evaluationsprozesse) sowie Zeit und Energie für Vertrauens- und Beziehungsaufbau/-arbeit stellen *Strategien* dar, wie diese Herausforderungen und Barrieren überwunden werden könnten (Welty Peachey & Cohen, 2016). Auch ein gemeinsames Verständnis davon, dass die Nutzung von Logischen Modellen für die Wirkung und Nachhaltigkeit von Programmen wichtig ist, würde nach Ansicht einiger Autor*innen dazu beitragen,

²⁶⁵ Bspw. die Verknüpfung der SfE-Landschaft mit politischen Prozessen, mangelndes Engagement und Ressourcen, die Aushandlung abweichender Zielsetzungen und Vorstellungen.

²⁶⁶ Hervorhebung durch die Autorin.

die Lücke zwischen Wissenschaft und Praxis zu schließen (Shin et al., 2020; Welty Peachey et al., 2019).

Nach Collison et al. (2016) spielen *Netzwerke*, die sich rund um SfE-Forschung gebildet haben, eine signifikante Rolle bei der Formierung und Konzipierung des globalen SfE-Sektors: „In this sense, research and researchers demonstrate mediating roles in organising and connecting the field of SDP“ (ebd., S. 418). Diese wichtige Position schürt Rollenerwartungen. So besteht die Möglichkeit, dass NGOs die Ankunft von internationalen Wissenschaftler*innen erwarten, deren Expertise die NGO-Arbeit in einem positiven Sinne bestätigen soll, besonders, wenn der*die Expert*in mit den Geldgebern in Verbindung steht. In diesem Fall äußern sich lokale Mitarbeitende möglicherweise aus Angst vor einem Projektende und Arbeitslosigkeit wenig kritisch bzgl. der eigenen Projektarbeit und bei Datenerhebungen besteht die Gefahr von Bias (ebd.; Jeanes & Lindsey, 2014). In diesem Zusammenhang spielt auch die *Identität* der Forschenden eine wichtige Rolle. Collison et al. (2016, S. 422) sprechen i.R. von eigenen Forschungserfahrungen von den negativen und positiven Auswirkungen ‚weiß und international‘ wahrgenommen zu werden: „[...] identifying or being identified through the lenses of whiteness and internationalism has had both positive and negative implications; positive in being viewed as experts, negative in being understood as separate from or privileged in relation to local people“. Darnell et al. (2018a) plädieren für Forschungspartnerschaften, die auf mehr als nur den individuellen Identifikationsprozessen basieren (wie Globaler Norden, weiblich, bürgerlich etc.) und zentrale *Überlegungen zu Machtstrukturen und Unterschieden* sowie deren Auswirkungen auf individuelle und institutionelle Beziehungen einbeziehen und reflektieren.

Des Weiteren sehen sich (Nachwuchs-)Wissenschaftler*innen Hindernissen ausgesetzt, sich im Forschungsfeld SfE zu engagieren. Neben einer bislang als problematisch wahrgenommenen Zusammenarbeit zwischen Wissenschaftler*innen und Praktizierenden stellt das *Hochschulsystem* selbst eine Hürde dar. SfE wird als Forschungsfeld oft nicht anerkannt. Zudem fehlen Forschungsgelder und der Druck, in kurzen Zeitabständen Artikel in hochrangigen Fachzeitschriften zu veröffentlichen, ist groß. Letzteres ist mit ei-

ner langjährigeren wissenschaftlichen Begleitung und Beratung von SfE-Projekten, Programmen oder Organisationen oft nicht möglich (Welty Peachey & Cohen, 2015). In Welty Peacheys und Cohens Studie (2019) werden allerdings mehrheitlich Wissenschaftler*innen aus dem Globalen Norden befragt (Darnell et al., 2018a). Die Schwierigkeiten, mit denen sich Forschende aus dem Globalen Süden auseinandersetzen müssen, um im Forschungsfeld Fuß fassen zu können, sind möglicherweise noch andere.

5.3 Erste Entwicklungen eines Forschungsfeldes

5.3.1 Das Forschungsfeld SfE – ein Überblick

Seit der Millenniumswende stößt der Sport bei staatlichen und nichtstaatlichen, nationalen und internationalen Akteur*innen zunehmend auf Akzeptanz, sowohl als eigenständiges Ziel als auch als Medium zur Erreichung von Entwicklungszielen. SfE erfährt eine zunehmende Institutionalisierung in der internationalen Zusammenarbeit und im globalen Entwicklungssektor, gestützt durch und flankiert von internationalen Organisationen wie den Vereinten Nationen und dem Commonwealth Sekretariat (Giulianotti, Coalter, Collison & Darnell, 2019). Der Umstand, dass das Arbeitsfeld SfE aus der Praxis herausgewachsen ist und theoretische Zugänge anfangs limitiert waren, begründet ein verzögertes Einsetzen des Forschungsinteresses (Darnell, 2012; Maguire, 2013). Dennoch tauchen Fragen auf, die das Interesse der Wissenschaft wecken:

„When and how can sport be a meaningful instrument for development or a development goal in itself? What are the possible outcomes of sport in specific situations for specific development goals? Which conditions are necessary to achieve these positive outcomes? What are the limits of sport for development programs? How can sport for development organizations adequately use the knowledge available in their program design and execution? How should they organize themselves, and with whom should they cooperate?“ (van Eekeren et al., 2013, S. 7)

Forschende wollen mehr über dieses ‚neue‘²⁶⁷ soziale Phänomen wissen und die Ansprüche und Forderungen („*Claims*“) erforschen, die u.a. seitens der politischen Akteur*innen hinsichtlich der Wirkungen von SfE gestellt werden (van Eekeren et al.,

²⁶⁷ Als ‚neu‘ wird hierbei nicht der Ansatz an sich, Sport als Medium zur persönlichen, sozialen und gesellschaftlichen Entwicklung zu nutzen, bezeichnet, sondern die gezielte Nutzung von Sport i.R. der EZ (vgl. Kidd, 2008; Levermore, 2008).

2013). Auch NGOs „[...] do more, rather than less, research“, um ihre Interessen, Fundraising und Programmgestaltung mit evidenzbasierten Ergebnissen untermauern zu können (Cronin, 2011, S. 15). Waren es Anfang der 2000er noch ‚eine Handvoll‘ an wissenschaftlichen Publikationen, stiegen die Beiträge in Wissenschaftsmagazinen ab ca. 2008 deutlich an (Schulenkorf, 2017; Schulenkorf et al., 2016). Der Zustand von nur wenigen und z.T. mangelhaften Evaluationen bei zeitgleicher Expansion des SfE-Feldes bis ca. Mitte der Nullerjahre wird von Levermore (2011, S. 339) als „Vakuum“ bezeichnet; Evaluationen waren i.R. der Projektextensionen ein „after-thought“ (ebd., S. 341).²⁶⁸ Allein im Jahr 2013 belief sich die Publikationszahl dann bereits auf über 100 Artikel (Schulenkorf, 2017; Schulenkorf et al., 2016).

Nicht nur die *Quantität* hat sich gesteigert, auch die *Qualität* der SfE-Studien hat sich im Laufe der Zeit verbessert: „As SFD has started to mature, studies have become more focused, differentiated, and critical in their approach; this, in turn, has allowed for a more nuanced design, implementation and evaluation of SFD as well as improved SFD practices, theory and policymaking“ (Schulenkorf, 2017, S. 244; vgl. auch van Eekeren et al., 2013, S. 34). Diesen willkommenen Veränderungen folgte eine Periode von „[...] uncritical ‚evangelical‘ accounts and assumed myopic powers of sport in the absence of robust evidence“ (Burnett, 2015, S. 386), was von vielen Forschenden verurteilt wurde (u.a. Coalter, 2013a; Darnell & Hayhurst, 2011; Darnell, 2012; Kay, 2012; Levermore & Beacom, 2008; Sugden, 2010). Schulenkorf (2017, S. 244) ist der Ansicht, dass die SfE-Community es nach 15 Jahren geschafft hat, diese „Krankheit zu heilen“, indem sie *Modalverben* (*can, could, may* etc.) verwendet, die die Möglichkeit oder Wahrscheinlichkeit unterstreichen, dass Sport einen Beitrag zu bestimmten Entwicklungszielen leisten *könnte* oder unter bestimmten Voraussetzungen *kann*.²⁶⁹

Ab Ende der Nullerjahre haben sich *Wissenschaftler*innen aus den verschiedensten Disziplinen* zunehmend mit dem Beitrag von Sport zu Entwicklungsprozessen befasst, häufig auch in Form von Berater*innen unter Vertrag von EZ-Agenturen (vgl. Burnett, 2015,

²⁶⁸ Neben fehlenden groß angelegten Evaluationen sieht er auch die Abkehr von Einzelprojekten in der internationalen EZ als einen Grund, dass EZ-Organisationen bis 2004 noch wenig Interesse zeigten (vgl. Levermore, 2011, S. 340).

²⁶⁹ S. weitere Ausführungen unter Kap. 5.4.4.

S. 387): „[The] proliferation of SDP organizations and SDP policy from non- governmental organizations (NGOs) and government entities has also captured the attention of scholars from across various social science disciplines, such as sociology, development studies, business management, and sport psychology“ (Welty Peachey, 2019, S. 241). Zur Erforschung wurden von Beginn an *zahlreiche theoretische Fundierungen, Ansätze, Designs und Methoden* genutzt.²⁷⁰ Entstandene Trends blieben jedoch meist isoliert in einzelnen Untersuchungsfeldern zurück (Massey, Whitley, Blom & Gerstein, 2019, S. 175f.) oder haben sich zu vereinzelt SfE-spezifische Theorien (weiter-)entwickelt. Das Forschungsfeld SfE wurde und wird isoliert von Wissenschaftler*innen unterschiedlicher Disziplinen betrachtet, zudem waren/sind an Evaluationsprozessen unterschiedliche Interessensgruppen aus dem sport-spezifischen und „non-sporting“ Sektor involviert.²⁷¹ Diese *Diversität an Betrachtenden, Betrachteten, Interessierten und Beeinflussenden* bedingt u.a. auch die Vielzahl an Studien, Publikationen und genutzten Ansätzen im Forschungsfeld SfE (ebd., S. 176; van Eekeren et al., 2013). Im Zusammenhang mit der theoretischen Fundierung des Forschungsfeldes betonen Levermore und Beacom (2012, S. 125) die Wichtigkeit „[...] of engaging with the established mainstream development discourse that provides an extensive body of theory [...]“, u.a. im Bereich von Wirkungsmessungen. Die Autor*innen sind der Ansicht, „Sport-for-development is an activity, *not an academic discipline*, nor indeed should it aspire to be“ (ebd., S. 134).²⁷²

Coalter (2010; 2013a) sieht das Forschungsfeld SfE mit drei Hürden konfrontiert: Mit der omnipräsenten und prägnanten Aussage „limited focus programmes and broad gauge problems“ (vgl. ebd., 2010, S. 295; 2013a, S. 18)²⁷³ beschreibt er den Versuch, „[...]

²⁷⁰ S. weitere Ausführungen unter Kap. 5.4.2 und Kap. 5.4.3.

²⁷¹ Aus dem sport-spezifischen Sektor „e.g. practitioners, academics, funders, policy- makers, governmental bodies, multilateral institutions, corporations, international non-governmental organizations“, aus dem non-sporting Bereich “e.g. non-profit management, international development, peacebuilding efforts, corporate social responsibility“ (Massey et al., 2019, S. 176).

²⁷² Hervorhebungen durch die Autorin.

²⁷³ Durch zahlreiche Beiträge hat Coalter Bezeichnungen und Aussagen wie „sport-for-development evangelists“ (ebd., 2013a, S. 20f.), „pessimism of the intellect, optimism of the will“ nach Gramsci (vgl. ebd., 2013a, S. 3ff.) und „limited focus programmes and broad gauge problems“ (vgl. ebd., 2010, S. 295; 2013a, S. 18) geprägt, auf den SfE-Kontext übertragen und den Diskurs beeinflusst (vgl. bspw. Burnett, 2015, S. 387). Während mit ‚SfE-Evangelist*innen‘ Lobbyist*innen und Enthusiast*innen gemeint sind, die unkritisch und ‚dogmatisch‘ an den positiven Nutzen von Sport glauben und diesen als ‚Allerheilmittel‘ oder ‚Allzweckwaffe‘ betrachten, beschreibt zweites Zitat das „intellektuelle und emotionale“ Spannungsfeld zwischen akademischen Vorbehalten und der Überzeugung, generiert aus Praxiserfahrungen meint (Colyer, 2013a, S. 5).

„Breitspurprobleme“ mittels „Schmalspurmaßnahmen“ zu lösen [...]“ (Brand, 2020, S. 111). Hiermit ist der Prozess von Politik-Lobbying gemeint, der über *unklar definierte Vorhaben auf Mikro-Ebene* versucht, *schwer nachvollziehbare Wirkungen auf Makro-Ebene* zu erzielen (vgl. Coalter, 2013a, S. 45).²⁷⁴ Als weitere Hürde sieht er die „relative Gleichgültigkeit“ von Politiker*innen hinsichtlich der Evaluation von SfE, die dieser hauptsächlich eine *Legitimitätsfunktion* zusprechen, d.h. Evaluation soll in erster Linie bereits getroffene politische Entscheidungen belegen. Als letztes nennt er die *methodischen Schwierigkeiten*, die mit der Messung von Wirkungen allgemein verbunden sind und im Speziellen, wenn Wirkungen alleinig dem Sport zugeschrieben werden sollen (ebd.).²⁷⁵

Gegenwärtig identifiziert Gadais (2019) *vier Hauptarten von Forschungsarbeiten* zum Thema SfE: (1) makrosoziologische Studien über die positiven Eigenschaften von SfE, (2) explorative Feld- und Fallstudien, (3) Studien über das Management und die Evaluation von SfE-Programmen und (4) Literature Reviews zu SfE.

5.3.2 Kontext und Meilensteine

Auch wenn Sport bereits im Zuge der *MDGs* in der Programmlinie *Sport, Entwicklung und Frieden* der Vereinten Nationen international Beachtung fand, wurde die Notwendigkeit und Dringlichkeit der umfangreicheren Evaluation von SfE-Programmen auch im Hinblick ihres Beitrags zu den *MDGs* erst im *Internationalen Jahr des Sports und der Leibeserziehung* 2005 deutlich und im Abschlussbericht zum Jahr mehrfach erwähnt (vgl. UNOSDP, 2006). Im selben Jahr wurde auch der erste internationale M&E-Workshop zu SfE-Programmen von UNICEF initiiert, an dem „key scholars, politicians, development agencies and practitioners“ teilnahmen (Burnett, 2015, S. 387). Ergebnisse des Work-

²⁷⁴ Da der politische Prozess von Überzeugungsarbeit und Verhandlungen geprägt ist, sind nach Weiss (1993, S. 96) solche „überhöhten Versprechungen“ hinsichtlich der Zielsetzungen üblich, um bspw. ein Programm bewilligt zu bekommen. Coalter (2013a, S. 19) spricht in diesem Zusammenhang auch von „displacement of scope“.

²⁷⁵ S. weitere Ausführungen unter Kap. 5.4.4.

shops waren u.a. zahlreiche Empfehlungen zur weiteren Vorgehensweise, eine Sammlung von *Best Practice* Beispielen und Tools beteiligter Akteur*innen sowie das „Sport in Development – A Monitoring and Evaluation Manual“ (vgl. Coalter, 2006).²⁷⁶

Die Betonung des Sports als wirksames und kosteneffizientes Instrument zur Erreichung entwicklungspolitischer Ziele i.R. politischer Agenden sowie diverse Institutionalierungsprozesse im UN-Kontext²⁷⁷ führten in den Nullerjahren zu ‚neuen Teamaufstellungen‘, ‚neuen Spielwiesen‘ und ‚neuen Spieler*innen auf dem Feld‘: Internationale Unternehmen kooperierten i.R. ihrer *Corporate Social Responsibility (CSR)*-Maßnahmen mit lokalen SfE-Initiativen (als ‚neue Teamaufstellungen‘); auch internationale Sportorganisationen (als ‚neue Spieler*innen‘) und nationale EZ-Agenturen nutzten Sportgroßereignisse (als ‚neue Spielwiesen‘) für CSR-Initiativen und EZ-Maßnahmen (Burnett, 2015).²⁷⁸

Die MDGs und später die SDGs zählen hierbei zu den wichtigsten Referenzpunkten von Fürsprecher*innen des Sports zur Erreichung entwicklungspolitischer Ziele (Dudfield, 2019). Wichtige Meilensteine zwischen den ‚großen Agenden‘ wie die Paris Deklaration oder die Accra Agenda for Action werden in der SfE-Literatur nur wenig erwähnt, auch wenn sie Auswirkungen auf die gesamte Evaluations-Praxis in der EZ hatten, auch auf jene mit Sportbezug (vgl. Preti, 2012, S. 309).²⁷⁹ Der Versuch, eine direkte Verknüpfung zwischen Sport und bestimmten Entwicklungszielen herzustellen und messbar zu machen, wurde zu Zeiten der MDGs noch weniger praktiziert als dies dann der Fall mit der Einführung der SDGs war und heute noch ist. In Anlehnung an die erste Review der *Sport*

²⁷⁶ Einen Überblick zur Entstehung und Inhalten des Manuals gibt Coalter (2009, S. 58ff.).

²⁷⁷ Gründung des *United Nation Office on Sport for Development and Peace (UNOSDP)* durch Kofi Annan, Adolf Orgi als Sonderbeauftragter für Sport, Entwicklung und Frieden, Einrichtungen von UNOSDP-Büros in Genf und New York (2001); Etablierung der Arbeitsgruppe *SDP IWG (Sport for Development and Peace International Working Group)* (2002); die SDP IWG versucht über ihren Bericht „Sport für Entwicklung und Frieden: Auf dem Weg zur Erreichung der Millennium-Entwicklungsziele“ SfE im UN-System zu operationalisieren (2003); UN-Resolution A/RES/58/5 beschreibt „Sport als zentrales, strategisches Aktionsfeld zur Erreichung der MDGs (UN General Assembly, 2003); Internationales Jahr der Erziehung durch Sport (2005) (vgl. Brand, 2020, S. 106f.).

²⁷⁸ Bspw. das *International Inspiration Programme* der FIFA 2012 i.R. der olympischen Spiele in London; das *YDF-Programm* der GIZ anlässlich der Fußballweltmeisterschaft 2010 in Südafrika (Burnett, 2013a, 2013b, 2015).

²⁷⁹ Allein der Paradigmenwechsel von einer Input-Output-Orientierung hin zu Outcomes und Impacts, der durch internationale Konsense wie den MDGs und der Paris Deklaration angetrieben wurde, setzte NGOs zunehmend unter Druck, ihre Arbeit zu evaluieren und Wirkungen zu messen (vgl. Preti, 2012, S. 309). Levermore und Beacom (2012, S. 133) erwähnen in diesem Zusammenhang, dass die internationalen Beschlüsse auch die Arbeit von großen SfE-Akteur*innen wie z.B. Right to Play beeinflusste.

for Development and Peace International Working Group (SDP IWG, 2007) wurde im Nachgang vielmehr von „thematischen Bereichen“ (RTP, 2008, S. viii) oder „Politikbereichen“ (Kay & Dudfield, 2013, S. 2) gesprochen, zu denen Sport einen Beitrag leisten kann.²⁸⁰

Im Rahmen der Präambel der Agenda 2030 wird Sport von Anfang an explizit als „important enabler of sustainable development“ betitelt (UN General Assembly, 2015a, S. 10)²⁸¹ und im weiteren Verlauf seitens internationaler Akteur*innen als „Allzweckwaffe‘ zur Erreichung der SDGs“ (Brand, 2020, S. 113) deklariert, wie bspw. des *SDG Funds* der Vereinten Nationen: „[...] sports is a tool that amalgamates the 2030 Agenda’s developmental outcomes with ease. From the local to the global level, examples of the impact of sports on unifying people and birthing development where most other strategies have failed, abound“ (SDG Fund, 2018, S. 7). Der Druck, einen sportspezifischen Beitrag zu den Entwicklungszielen herausstellen zu können, wird dementsprechend erhöht, auch wenn Sport in keinem weiteren SDG oder Unterziel erwähnt wird (vgl. UNESCO, 2017, S. 26). Der Ruf nach „measured, evidence-based approaches“ (Dudfield & Dingwall-Smith, 2015, S. 85) verdichtet sich im ‚Sport und SDG-Diskurs‘, die Notwendigkeit von robusten M&E-Systemen wird dringlicher. Die seitens des Commonwealth initiierte Studie (als ‚analysis‘ betitelt) „Sport for Development and Peace and the 2030 Agenda for Sustainable Development“ identifizierte sechs SDGs „[...] where sport-based approaches could make effective and cost-efficient contributions“ (ebd., S. 3).²⁸² Die Studie war wegbereitend für weitere Schritte, bspw. bei der sechsten internationalen

²⁸⁰I.R. von Publikationen (SDP IWG, 2007; RTP, 2008; Kay & Dudfield, 2013; Lindsey, Kay, Jeanes & Banda, 2017) werden nur teilweise alle MDGs aufgeführt und ein möglicher Sportbeitrag beschrieben (vgl. RTP, 2008, S. 11f.). In einer der ersten SfE-Reviews (vgl. SDP IWG, 2007) werden die MDGs zwar als Referenz angeführt, allerdings wird auf den grundsätzlich wichtigeren Bezug von Sport zu den Menschenrechten verwiesen (vgl. ebd., S. 9). In einer Veröffentlichung von *Right to Play* (RTP) als Teil der SDP IWG (RTP, 2008, S. 11) wird der Beitrag zu jedem MDG explizit beschrieben und auf die Notwendigkeit eines M&E-Frameworks i.R. der Entwicklung effektiver Policies und Programme hingewiesen (vgl. ebd., S. 270). In Anlehnung an die erste Review wird in Ausarbeitungen des Commonwealth Sekretariats (Kay & Dudfield, 2013) der Beitrag von Sport den folgenden Bereichen zugeordnet, mit denen verschiedene MDGs verknüpft sind: „Youth, Health, Education, Gender, Equity and Inclusion, Peace and Stability“ (ebd., S. 65).

²⁸¹ „This was the first time that any overarching policy for global development included such a wide-ranging statement on sport and, as such, it represented recognition of the significant expansion in activity and advocacy associated with SDP since the turn of the millennium“ (Lindsey & Darby, 2019, S. 794).

²⁸² Aus der Veröffentlichung geht nicht genau hervor, welches Studiendesign bzw. welche methodische Vorgehensweise gewählt wurde, außer das neben „desk research“, der Konsultierung von Expert*innen ein Analyseframework erarbeitet wurde (vgl. Dudfield & Dingwall-Smith, 2015, S. 6).

Sportministerkonferenz im Jahr 2017 (*International Conference of Ministers and Senior Officials Responsible for Physical Education and Sport - MINEPS VI*): Da die offizielle Liste an SDG-Indikatoren über den Wirkungsbereich von SfE-Initiativen hinausgeht (Dudfield, 2019), wurde i.R. des *Kazan Action Plans* der *MINEPS VI* die Entwicklung gemeinsamer Indikatoren beschlossen, „for measuring the contribution of physical education, physical activity and sport to prioritized SDGs and targets“ (UNESCO, 2017, S. 23).²⁸³ Der Beschluss nimmt sich damit sowohl des international aufgebauten Drucks als auch der diagnostizierten Gefahr an und sieht sowohl die Erhebung von aggregierbaren Daten zu den allgemeinen SDG-Indikatoren auf nationalem Level (vgl. Commonwealth Secretariat, Juni 2018, S. 7) als auch die Entwicklung von sportspezifischen „Modell-Indikatoren“ zu ausgewählten SDGs vor (vgl. ebd., S. 8).

5.3.3 Institutionalierungs- und Professionalisierungsprozesse

Die diversen Entwicklungen hinsichtlich der Institutionalisierung und Professionalisierung im Forschungs- und Evaluationsbereich von SfE umfassend zu überblicken, stellt eine Herausforderung dar. Dass die Vereinten Nationen hierbei von Anfang an eine führende Rolle spielten, ist unumstritten: Die Gründung des *United Nations Office on Sport for Development and Peace (UNOSDP)* 2001 und der *SDP IWG* ab 2005, an denen zahlreiche staatliche und nichtstaatliche Akteur*innen auch aus der Wissenschaft beteiligt waren, waren wegbereitend für wichtige Studien (z.B. *SDP IWG*, 2008), Standards (z.B. Coalter, 2006) und Konferenzen. 2003 fand die erste internationale *Next Step Conference* in Holland statt, an der Expert*innen und Praktiker*innen aus 45 Ländern teilnahmen (RTP, 2008; van Eekeren et al., 2013, S. 31f.). Das Format wurde bis 2014 in verschiedenen Ländern durchgeführt (2005 Sambia, 2007 Namibia, 2011 Trinidad und Tobago, 2014 Indien). Vergleichbare Konferenzen fanden im weiteren Verlauf nicht statt (vgl. sportanddev.com, o.D.). Im wissenschaftlichen Bereich wird SfE i.R. unterschiedli-

²⁸³ Der *Kazan Action Plan* und der *UN Action Plan on Sport for Development and Peace* (UN General Assembly, 2018) werden in der SfE-Literatur als wichtige Policy Dokumente hervorgehoben (vgl. Whitley et al., 2020, S. 21/24), die die Wichtigkeit von M&E sowie die Entwicklung von international abgestimmten Systemen, „standards, indicators and benchmarks“ betonen (UN General Assembly, 2018, S. 5).

cher Konferenzen und Symposien platziert, bspw. im Bereich Sportsoziologie, -management, -politik oder der Entwicklungsforschung.²⁸⁴ Auch wenn die Arbeit des UNOSDP ob einer fehlenden Strategie und eines fehlenden Einflusses kritisiert wurde, hinterließ die Schließung des Büros im Jahr 2017 ein „transnational leadership vacuum“ (Giulianotti et al., 2019, S. 422).²⁸⁵ Das *Commonwealth Sekretariat* hat eine führende Rolle bei der Weiterentwicklung von Evaluationsstandards zur Messung des Beitrags von Sport zu den SDGs übernommen und wird hierbei von namenhaften SfE-Forschenden und Expert*innen von Beratungsfirmen wie *inFocus* unterstützt (vgl. Dudfield & Dingwall-Smith, 2015; Lindsey & Chapman, 2017; Commonwealth Secretariat, 2019).

inFocus hat sich als Beratungsfirma u.a. auf die Evaluation von SfE in der EZ spezialisiert. Zu ihren Klient*innen gehören SfE-Akteur*innen, u.a. die GIZ, der internationale Basketballverband (FIBA) und mehrere SfE-NGOs.²⁸⁶ In Zusammenarbeit mit der *Laureus Sport for Good Stiftung*, dem Netzwerk *streetfootballworld* wurden *M&E Qualitätsstandards für SfE-Organisationen* sowie der Leitfaden „7 Steps to effective impact measurement“ entwickelt (vgl. *inFocus*, 2014, 2016). Die Firma bietet kostenpflichtige Softwares und Schulungen über Web-Seminare und E-Learning an. Auch bei der Entwicklung eines Toolkits zur Messung des Beitrags von Sport zu den SDGs über Modellindikatoren waren Beratende der Firma beteiligt (vgl. Commonwealth Secretariat, 2019).²⁸⁷ Die deutsche Beratungsfirma *PHINEO* hat 2015 ebenfalls eine Studie zur Wirksamkeit von Sportinitiativen publiziert, u.a. in Kooperation mit dem DOSB (vgl. Kubek & von der Ahe, 2015).²⁸⁸

²⁸⁴ 2016 fand ein *SDP Research Symposium* in Atlanta statt, mit dem Ziel, SfE mit der Entwicklungsforschung zu verbinden (vgl. Chawansky et al., 2017).

²⁸⁵ Das UNOSDP sollte durch eine direkte Partnerschaft zwischen UN und dem *Internationalen Olympischen Komitee (IOC)* ersetzt werden (Giulianotti et al., 2019). Giulianotti et al. (2019, S. 422f.) empfehlen drei „overlapping, yet imperfect, scenarios“, die benanntes Vakuum füllen könnten: (1) IOC und das Commonwealth Sekretariat bauen ihre bereits führenden Rollen aus; (2) Unterschiedliche multilaterale Partnerschaften und Netzwerke nehmen sich speziellen Themen an, bilden sog. „transnational advocacy networks“ (TANs) und setzten SfE-Programme gemeinschaftlich in bestimmten Regionen um, die von UNICEF Länderbüros geleitet werden könnten; (3) eine vermarktete Variante von SfE mit Verbindung zu CSR-Initiativen von Unternehmen.

²⁸⁶ U.a. Fight for Peace (UK, Brasilien), Promundo (Brasilien), Mathare Youth Sports Association (Kenia) und Magic Bus (Indien).

²⁸⁷ *inFocus* wurde 2007 von Damian Hatton gegründet, der zuvor *Street League*, ein soziales SfE-Programm in England ins Leben gerufen hatte. Die Laureus Sport for Good Stiftung und *streetfootballworld* sind von Anfang an strategische Partner der Consultancy (*inFocus*, 2021, 10. Juni).

²⁸⁸ I.R. der Studie wurden unterschiedliche sportpädagogische Initiativen in Deutschland untersucht und mit einem sog. *Wirkt-Siegel* ausgezeichnet (vgl. Kubek & von der Ahe, 2015).

Im Nachgang des ersten UNICEF-Workshops zu Monitoring und Evaluation im Jahr 2005 initiierten verschiedene *Zusammenschlüsse von Wissenschaftler*innen* gemeinsame Foren oder Workshops und beteiligten sich an Artikeln zu Sonderausgaben von Wissenschaftsmagazinen und Kapiteln zu Sammelbänden. Resultierend aus der Bündelung verschiedener Wissenschaftsströme sowie Partnerschaften aus Wissenschaftler*innen, Praktizierenden und (politischen) Geldgebern entstanden weitere Sub-Netzwerke in bestehenden Netzwerken (Burnett, 2015). Auch die Zunahme an Beratenden, die EZ-Organisationen wissenschaftlich begleiten (extern), bedingte eine immer weitreichendere Vernetzung der Expert*innen (ebd., Levermore, 2011). Laut Burnett (2015, S. 388) ist die SfE-Wissenschaftscommunity gegenwärtig nicht durch eine geografische Nähe der regionalen Zusammenarbeit, sondern hauptsächlich durch eine *Diskurszugehörigkeit* geprägt. In diesem Sinne stellt sie eine virtuelle und globale Gemeinschaft dar, die vom akademischen (Ausbildungs-)Hintergrund, disziplinärer Zugehörigkeit und der ideologischen Positionierung beeinflusst und häufig vom Tätigkeitsumfeld und der eigenen Praxis inspiriert ist.

In vielen Beiträgen wird nach wie vor die *mangelnde Transparenz*, d.h. SfE-Organisationen veröffentlichen nur wenig Evaluationsergebnisse, die *einseitige Berichterstattung*, d.h. die *Fokussierung auf positive Ergebnisse*, sowie der *schlechte Zugang zu und die nutzerunfreundliche Aufbereitung von Ergebnissen* aus Forschung und Evaluation kritisiert (vgl. Engelhardt, 2019, S. 128; Cronin, 2011, S. 39; Darnell et al., 2018a, S. 144; 2019, S. 14; Levermore, 2011, S. 342; Schulenkorf et al., 2016, S. 35; Welty Peachey & Cohen, 2015, S. 20; Whitley, Massey, Camiré, Blom, Chawansky & Forde, 2019a, S. 181; Whitley et al., 2019b, S. 6f.). Es ist von einer „sizeable gap between information and access“ die Rede, die es zu schließen gilt (Whitley et al., 2019b, S. 8). Die *International Platform on Sport and Development sportanddev.org* hat sich in diesem Zusammenhang zu einem unabdingbaren Informationsmedium entwickelt: Neben der Vernetzung von Akteur*innen, der Abbildung von jeglichen SfE-Neuigkeiten und der Ankündigung von Events (bspw. Konferenzen) werden hier auch Studien und Evaluationsberichte zur Verfügung gestellt. Vor allem für Nicht-Akademiker*innen stellt die Plattform einen wichtigen Informationszugang dar, weshalb er zu weiteren wissenschaftlichen Beiträgen weiter ausgebaut werden sollte (vgl. Whitley et al., 2019b, S. 6). Zudem wurde 2013 das frei

zugängliche Onlinejournal *Journal of Sport for Development (JSFD)* mit der Mission veröffentlicht, „[...] to advance, examine and disseminate evidence and best practices for programmes and interventions that use sport to promote development, health and/or peace“ (jsfd.org, o.D.).

Der *Mangel an lokalen Kapazitäten*, d.h. lokales Personal oder Forschende, die SfE-Studien durchführen und/oder M&E-Systeme in ihren Organisationen nachhaltig verankern, spielt in diesem Zusammenhang ebenfalls eine große Rolle (vgl. Cronin, 2011, S. 39; Kay, 2012, S. 892; Preti, 2012, S. 313). Laut Whitley et al. (2019b, S. 5) besteht seitens lokaler Organisationsmitarbeitenden zwar ein großes Interesse am Thema, allerdings sind nur wenige ausgebildet und mit M&E-Aufgaben überfordert. Die gezielte Förderung lokaler M&E-Kapazitäten in Organisationen und an lokalen Universitäten sowie einer Online-Plattform, die „improved, accessible training and support“ anbieten, könnte Abhilfe schaffen, (vgl. ebd., S. 7ff.). Einzelne Akteur*innen wie die Laureus Sport for Good Stiftung oder streetfootballworld etablierten sog. *Learning Communities* oder bieten Foren an, in denen Expert*innen Wissen teilen und sich Praktizierende austauschen (Whitley et al., 2020). *Sportanddev.org* bietet seit 2020 einen *Massive Open Online Course (MOOC)*, der in Kooperation mit der australischen Regierung und dem Commonwealth Sekretariat erarbeitet wurde und u.a. Inhalte zu M&E und Forschung beinhaltet (ebd.; Sanders, 2020, 25. August).

Lange wurde SfE mit einem „low standing academic status“ in Verbindung gebracht (van Eekeren et al., 2013, S. 31): „Studying sport for development was approached as a hobby gotten out of hand, and not as a factor contributing to a serious academic career.“ Nach wie vor ‚kämpft‘ SfE, u.a. aufgrund der *Multidisziplinarität*, um Anerkennung in verschiedenen Wissenschaftsbereichen (vgl. Levermore & Beacom, 2012, S. 15f.; Welty Peachey & Cohen, 2015, S. 20). Da auch Studierende ein wachsendes Interesse am Themenfeld zeigen, gewinnt SfE zunehmend an „Glaubwürdigkeit“ in „akademischen Kreisen“ (vgl. Welty Peachey & Cohen, 2015, S. 22) und SfE wird zunehmend an Hochschulen verankert, bspw. i.R. des internationalen Masterprogramms *International Sport Development*

and Politics an der DSHS.²⁸⁹ *Forschungsk Kooperationen* zwischen Organisationen, Wissenschaftler*innen, Forschungseinrichtungen und politischen Akteur*innen sind schon seit geraumer Zeit gefordert, erwünscht und nehmen zu (vgl. Donnelly, Atkinson, Boyle & Szto, 2011; Kay & Dudfield, 2013; Svensson, Hancock & Hums, 2016; Welty Peachey et al., 2019; Whitley et al., 2020),²⁹⁰ was sich bspw. am *Erasmus+* Projekt *GloKnoCo*, der *Global Sport for Development and Peace Knowledge Collaborative* zeigt, an der Universitäten und NGOs beteiligt sind (vgl. GloKnoCo, 2019).

5.3.4 Standardisierungsprozesse

Whitley et al. (2019b, S. 4) merken an, dass SfE „[...] has been critiqued about a lack of clarity, which begins with the term itself“. Unklarheiten in der Terminologie beziehen sich nicht nur auf den Gegenstand SfE selbst, sondern auch auf die Unterscheidung, Nutzung und Einsortierung der Begriffe *research*, *evaluation* und *M&E*, was die Ausführungen in Kapitel 5.1 bereits verdeutlichten. Demnach ist es nicht verwunderlich, dass sich nur wenige Handbücher auf allgemeingültige Evaluationsstandards beziehen (vgl. Cronin, 2011, S. 28). Dennoch existieren im SfE-Kontext *Manuale, Leitfäden und Beiträge mit dem Fokus auf M&E*: Nach der Veröffentlichung des ersten M&E-Manuals für den sportbezogenen EZ-Kontext (vgl. Coalter, 2006), folgten noch weitere Manuale, teilweise seitens unterschiedlicher Netzwerke und nichtstaatlicher Akteur*innen (bspw. NSD, 2009; Elsemann, Hebel, Jäger & Daraspe, o.J.; CARE, Mercy Corps & Schwery Consulting, 2008; inFocus 2014, 2016) (vgl. Cronin, 2011, S. 22ff.), sowie Beiträge i.R. von Sammelbänden (bspw. Engelhardt, 2019; Sherry, Schulenkorf & Philips, 2016). Über Homepages werden u.a. Definitionen, Methoden und Instrumente online zur Verfügung gestellt (bspw. sportanddev.org, 2022, 24. März; S4D Resource Toolkit, o.D.a), andere

²⁸⁹ S. weitere Bsp. bei Van Eekeren et al. (2013, S. 32) und Whitley et al. (2020, S. 26).

²⁹⁰ Der *Commonwealth Guide to Advancing Development through Sport* (Kay & Dudfield, 2013, S. 99) verankert Kooperationen mit Hochschulen i.R. des Indikatorenframeworks: „SDP research projects and initiatives backed by a commitment from the higher education sector. [...] While international research networks and partnerships play an important role, increasing the opportunities for academics and post-graduate students to conduct research within their own national contexts should be considered a priority.“

Toolkits sind der Öffentlichkeit nicht zugänglich (vgl. Cronin, 2011, S. 26). NGO-Vertretende sehen nach wie vor Bedarf an M&E-Leitfäden sowie standardisierter Instrumente, bspw. für den Bereich der Wirkungsmessung (vgl. Whitley et al., 2019b, S. 5).

Das Commonwealth Sekretariat hat im Sinne eines Standardisierungsprozesses in der vergangenen Dekade eine zunehmend wichtige Rolle eingenommen: Neben Beiträgen, die die Wirkung von Sport bzgl. bestimmter Entwicklungsziele näher bestimmen (vgl. Dudfield & Dingwall-Smith, 2015; Lindsey & Chapman, 2017), beinhaltet der „Guide to Advancing Development through Sport“ (Kay & Dudfield, 2013) ein konkretes Indikatorenset, das darauf abzielt, SfE durch „policy, strategy and support mechanisms“ zu stärken. Eine der sechs Zielsetzungen bezieht sich auf „Evidence-based Programming, monitoring and evaluation“.²⁹¹ Die Autor*innen betonen hierbei die damit verbundenen Herausforderungen:

„Monitoring and evaluating the contribution sport-based approaches make towards development objectives is challenging and requires a more developed methodology than tracking sport participation or involvement in programmes and projects. The development areas to which sport is best placed to contribute are complex domains that usually encompass multi-sector approaches. Sport will often be one of multiple interventions employed making it difficult to establish causal links between a sport-based intervention and stated development objectives. Monitoring and evaluation that recognises this complexity, and that is framed within broader evaluation systems and based on well-developed theories and models, are often the most appropriate“ (Kay & Dudfield, 2013, S. 101).

Wie die angesprochene Komplexität i.R. von M&E-Systemen bedacht werden kann, wie und welche übergeordneten Systeme und Frameworks sowie „gut-designed“ Theorien und Modelle genutzt werden könnten, bleibt i.R. der MDGs noch im Unklaren. Konkretere Vorschläge und Lösungswege werden sich erst mit der Einführung der SDGs in den Expertenkreisen entwickeln.

Die im Kazan Action Plan der MINEPS VI beschlossene *Entwicklung gemeinsamer Indikatoren* wurde im weiteren Verlauf durch das Commonwealth Sekretariat koordiniert. Die

²⁹¹ Das Indikatorenset beinhaltet sechs sog. *Principles*, die bspw. auch die Integration von Sport im allgemeinen Entwicklungssektor als wichtigen „policy, strategy and support mechanism“ beinhalten. Zur Zielsetzung „evidence-based programming and monitoring and evaluation“ gehören drei Indikatoren zur Kooperation und Unterstützung von Hochschulen (6.1), M&E-Maßnahmen zur Sportpartizipation (6.2) sowie zur Ermittlung des Beitrags von sportbezogenen Interventionen zu Entwicklungszielen (6.3) (vgl. Kay & Dudfield, 2013, S. 68f.).

aktuelle Veröffentlichung „Measuring the contribution of sport, physical education and physical activity to the Sustainable Development Goals - Toolkit and model indicators - Version 3.1“ (Commonwealth Secretariat, 2019) beinhaltet ein *Sport and SDG Measurement Framework* und dazugehörige *Input-, Outcome- und Impact-Indikatoren* zu drei Kategorien, wobei die Indikatoren z.T. aggregierbar sind (vgl. ebd., S. 40ff.).²⁹² Der gesamte Prozess, u.a. auch aktuelle Beratungs-, Test- und Validierungsprozesse zu den Indikatoren, wird von zahlreichen Stakeholdern, darunter auch nationale Regierungen und institutionelle Akteur*innen, begleitet (Sanders, 2019, 23. April).²⁹³ Sanders (2019, 23. April) fasst zusammen: „The need for established methods to measure the impact of sport on the achievement of the Sustainable Development Goals, and validating sport as an enabler of social and economic development is overdue. But progress is being made, and the data gap is closing.“

5.4 Schwerpunkte, theoretische Zugänge und Ansätze

5.4.1 Inhaltliche Schwerpunkte

Wie bereits geschildert, richtete sich der SfE-Kontext inhaltlich stark an internationalen Agenden und dazugehörigen Entwicklungszielen aus, d.h. an den MDGs und später den SDGs. In einer der ersten Reviews zu bereits bestehender SfE-Literatur (vgl. SDP IWG, 2007) werden *Best Practices*, Erkenntnisse und Forschungslücken zu den folgenden Themen aufgezeigt: „Child and youth development and education“, „Health objectives“, „Gender equity“, „Inclusion, health and well-being of people with disabilities“, „Social cohesion, prevent and reduce conflict and build peace“ (ebd., S. 3), die auch im weiteren Verlauf aufgegriffen werden (RTP, 2008; Kay & Dudfield, 2013). Van Eekeren et al. (2011,

²⁹² Und zwar der Kategorie 1 und 2. *Category 1: Common indicators* – a small suite of global indicators to be collected consistently across all countries (bspw. % of females who actively participate in sport, fitness and active recreation); *Category 2: Context-/SDG-specific indicators* – a wider set of indicators that could be collected at the regional, national or sub-national level to guide the measurement efforts of a broad range of stakeholders, based on context and aligned with regional or national development priorities and the SDGs (bspw. % of population engaging in moderate and/or vigorous physical activity through sport, fitness and active recreation); *Category 3: Programmatic indicators* – a common methodology for categorising programmatic or project-level interventions that produces results explicitly linked to prioritised SDGs and targets (bspw. # of young people aged 15-24 whose attitude toward getting a job has improved) (Commonwealth Secretariat, 2019, S. 40f.).

²⁹³ Zu detaillierten Ausführungen zur Einbettung des Prozesses und der Beteiligten s. Sanders (2019, 23. April) und Abbasi (2019, 25. November).

S. 3) sind der Ansicht, dass Sport nur einen direkten Beitrag zu *Gesundheitszielen* leisten kann:

„According to the academic literature, health is the only overarching outcome that shows a direct relationship with physical activity. For all other overarching development outcomes, it is hard to prove a direct causal relation between sport for development activities and development outcomes. It is important to approach sport as ‘site for socialization experiences’, not causes of socialization outcomes. It is not sport in itself that leads automatically to development. Sport takes place in a complex social context in which various factors influence the final outcomes of a sport for development program.”

Mit der Einführung der SDGs identifizierten Dudfield und Dingwall-Smith (2015) sechs SDGs, die besonders geeignet sind, einen wirksamen und effizienten Beitrag zu leisten.²⁹⁴ SDG 17 „Partnerschaften zur Erreichung der Ziele“ wurde zusätzlich als wichtiges Querschnittsziel positioniert, das für alle SfE-Akteur*innen relevant ist (vgl. ebd., S. 3). Die Studie wird von Lindsey und Chapman (2017) spezifiziert, indem je Ziel nach einem direkten und indirekten/verknüpften Beitrag („direct/linked contribution“) unterschieden wird und dazugehörige Unterziele der SDGs gelistet werden (vgl. ebd., xvii ff.). Die SDG-Modellindikatoren des Commonwealth Sekretariats (2019) beziehen sich sowohl auf diese Vorarbeiten als auch auf die i.R. des Kazan Action Plans bezeichneten *acht Impact Areas* (ebd., S. 44): (1) Sport for health and well-being for all (SDG 3); (2) Sport for making cities, inclusive, safe, resilient and sustainable (SDG 11); (3) Sport for quality education and lifelong learning for all (SDG 4); (4) Sport for peaceful, inclusive and equitable societies (SDG 10, SDG 16); (5) Sport for economic growth and full, productive employment (SDG 8); (6) Sport for gender equality and empowerment for all women and girls (SDG 5); (7) Sport for sustainable consumption and action against climate change (SDG 12, SDG 13); (8) Sport for effective, accountable and inclusive institutions (SDG 16).

Über die Arbeit der internationalen Organisationen hinaus identifizieren Giulianotti et al. (2019, S. 428ff.) die Bereiche „disability, social inequalities, environment, safeguarding, refugees, anti-radicalization, social enterprises, social capital“ als künftig wichtige Evaluationsgegenstände und -themen. Laut Schulenkorf et al. (2016, S. 22) beschäftigen

²⁹⁴ SDG 3 (Gesundheit und Wohlergehen), 4 (hochwertige Bildung), 5 (Geschlechtergleichheit), 8 (menschenwürdige Arbeit und Wirtschaftswachstum), 11 (nachhaltige Städte und Gemeinden) und 16 (Frieden, Gerechtigkeit und starke Institutionen).

sich die meisten Studien seit 2000 hauptsächlich mit „[...] social and educational outcomes related to youth sport and with football (soccer) as the most common activity“. Die Themen „social cohesion“, „education“ und „disadvantaged communities“ sind in vielen Studien zu finden, während andere wie „disability, gender equality, and livelihoods“, wozu bspw. auch Themen wie „employability“ gehören, bislang noch zu wenig repräsentiert sind (ebd., S. 33f.).

Die Zuordnung des Beitrags von Sports zu einzelnen SDGs und dazugehöriger Unterziele sowie die zeitgleiche Positionierung von Sport als „cross-cutting tool across different development agendas“ könnte als *paradox* bezeichnet werden (Lindsey & Darby, 2019, S. 796), spiegelt sich jedoch im ‚integrierten, unteilbaren und verflochtenen‘ Charakter der Agenda 2030 wider. Lindsey und Darby (2019) schlagen das Konzept der *Politikkohärenz*²⁹⁵ vor, um diese wichtige *Intersektionalität* von Sport zu unterschiedlichen SDGs und Zielen zu identifizieren und zu untersuchen, wie bspw. im Bildungs- und Gesundheitsbereich sowie im Hinblick auf Migration. Den Autor*innen nach könnte diese Betrachtungsweise eine größere Aufmerksamkeit und Akzeptanz in etablierten EZ-Sektoren sichern und Prioritäten für die weitere Entwicklung des Arbeitsfeldes aufzeigen (vgl. ebd., S. 807).

5.4.2 Zur Fundierung und ‚Theoretisierung‘ von SfE

Die *fehlende theoretische Fundierung* von Forschung im Bereich SfE wird von zahlreichen Autor*innen bereits früh angemahnt, die Forderung danach und der Ruf nach Anschlussfähigkeit an andere Disziplinen zunehmend dringlicher (vgl. Coalter, 2010, S. 295; Coalter, 2013b; Darnell & Dao, 2017, S. 26; Lyras & Welty Peachey, 2011, S. 311; Massey et al., 2019, S. 179; Welty Peachey & Cohen, 2015, S. 24; Zanotti & Stephenson, 2019, S. 172). Zwar verweisen Forschende hierbei auf bestehende Literatur aus anderen Disziplinen, die genutzt werden kann, allerdings schließt nach wie vor beinahe jede SfE-Studie mit dem Ruf nach theoretisch-fundierter Forschung (vgl. Biermann, 2016, S. 27).

²⁹⁵ *Politikkohärenz* wird in der Agenda 2030 i.R. des Ziele 17.3 und 17.4 ausdrücklich erwähnt (vgl. UN General Assembly 2015a, S. 29). Zudem wurden die OECD/DAC-Evaluationskriterien 2019 um das Kriterium der Kohärenz ergänzt (DEval, 2019, 17. Dezember) – mit dem Ziel „[...] to better capture linkages, systems thinking, partnership dynamics, and complexity“ (OECD, 2019, S. 3).

Trotz des diagnostizierten Mangels an theoretischer Fundierung und der Feststellungen, dass hierzu „noch viel Arbeit bevorsteht“ (vgl. Zanotti & Stephenson, 2019, S. 172), fand in der vergangenen Dekade eine immense Entwicklung statt (Massey et al., 2019; Schulenkorf, 2016). Nicht zuletzt die Beiträge zur gesellschaftswissenschaftlichen Rahmung im *Routledge Handbook of Sport for Development and Peace* (Collison et al., 2019) sowie die Metaanalyse von Welty Peachey et al. (2019) zu den theoretischen Fortschritten im Bereich SfE veranschaulichen, dass die Forschergemeinde sich der Dringlichkeit annimmt und versucht, sich einen Überblick über theoretische Zugänge zu verschaffen, sie zu sortieren und gezielter einsetzbar zu machen.

Verschiedenste theoretische Rahmungen aus den *Sozial- und Geisteswissenschaften* wurden zumeist i.R. von Fallstudien genutzt (vgl. Schulenkorf et al., 2016, S. 36; Levermore & Beacom, 2009).²⁹⁶ Die meisten Einflüsse und Anstöße kommen von Wissenschaftler*innen aus dem Bereich der *Soziologie*, die argumentieren, dass Forschung im Bereich SfE „[...] cannot be detached from the social and political conditions in the local context“ (Massey et al., 2019, S. 176). Zudem ist die sportbezogene EZ genauso von *hegemonischen Diskursen* bestimmt wie die EZ im Allgemeinen, besonders bzgl. der Bedeutung und der Praxis im Zusammenhang mit dem Begriff *Entwicklung*. Es wird angenommen, dass „one of the first and most basic challenges of the whole sport and development field has to do with the *conception of development* itself“ (Hartmann & Kwauk, 2011, S. 286; s. auch Darnell, 2014, S. 3). Demnach ist SfE mit denselben politischen Herausforderungen konfrontiert, die in der Entwicklungsforschung in den vergangenen 20 Jahren identifiziert wurden (vgl. Darnell & Black, 2011, S. 373).²⁹⁷ Vor dem Hintergrund der *Soziologie, der Politikwissenschaften und Entwicklungsforschung* nutzten Forschende theoretische Rahmungen u.a. zur Erörterung und Erklärung von (ungleichen) Machtverhältnissen, Governance-Strukturen, gesellschaftlichen Konstrukten, Systemen, Kontexten und Entwicklungen, darunter postkoloniale (z.B. Hayhurst, 2009; 2014; Darnell, 2007; 2014; Darnell & Hayhurst, 2012; Lindsey & Grattan, 2012; Tiessen,

²⁹⁶ Die folgenden Erläuterungen sollen einen Einblick zur *Vielfältigkeit genutzter Theorien* aus unterschiedlichen Disziplinen geben und exemplarische Studien benennen, sie erheben jedoch keinen Anspruch auf Vollständigkeit.

²⁹⁷ Levermore (2009, S. 26ff.) gibt einen Überblick zu Sichtweisen und Perspektiven in der internationalen Entwicklungsdebatte, die auch den Bereich SfE beeinflussen.

2011)²⁹⁸, feministische (bspw. Hayhurst, Giles & Radforth, 2015)²⁹⁹ und politisierende Theorien, u.a. zivilgesellschaftliche und Verteilungstheorien nach Putnam, Coleman und Bourdieu (Coalter, 2007a; 2013a; Giulianotti, Hognestad & Spaaij, 2016;), (Akteurs-)Netzwerktheorien (Burnett, 2009; Darnell et al., 2018b; Henne, 2017; Richelieu & Webb, 2019), Capability-Ansätze (Darnell & Dao, 2017), Community Capacity Ansätze (Edwards, 2015), anthropologische Ansätze und Theorien (bspw. Burnett, 2008; Collison & Marchessault, 2018)³⁰⁰, Globalisierungstheorien (Giulianotti & Armstrong, 2014), Foucaults Diskursanalyse (z.B. Darnell, 2007; Harris & Adams, 2016; Nicholls et al., 2011) sowie Konzepte des kritischen Realismus und Pragmatismus (bspw. Schulenkorf & Sugden, 2011; Sugden, 2010, 2014).

Forschende, die SfE aus einer *pädagogischen oder psychologischen Perspektive* betrachten, nutzen u.a. Ansätze der kritischen Pädagogik nach Freire und Giroux (z.B. Oxford & Spaaij, 2017; Spaaij & Jeanes, 2013), das Teaching Personal and Social Responsibility (TPSR) Modell (bspw. Mandigo, Corlett & Ticas, 2016; Wright, Jacobs, Ressler & Jung, 2016) sowie Entwicklungsmodelle wie der ökosystemische Ansatz nach Bronfenbrenner (z.B. Biermann, 2016) und der positiven Jugendentwicklung (eng.: positive youth development - PYD) (z.B. Holt et al., 2017)³⁰¹, Selbstbestimmungs- und Selbstwirksamkeitstheorien u.a. nach Bandura (bspw. Coalter & Taylor, 2010; Farello, Blom, Mulvihill & Erickson, 2019; Kaufman et al., 2016) oder Allports Kontakthypothese (bspw. Edwards et al., 2015; Tuohey & Cognato, 2011). Die Ansätze versuchen, u.a. den Beitrag von Sport zur Förderung positiver Entwicklung von Jugendlichen, von Life Skills oder zur Friedensförderung und dem Abbau von Vorurteilen zu untersuchen und Lernzusammenhänge bspw. im Gesundheitsbereich zu erklären (Guest, 2019). Haudenhuyse, Theeboom und Nols (2013) sowie Thorpe (2016) machen in diesem Zusammenhang auf die ‚passive‘

²⁹⁸ S. zum Überblick Levermore (2009, S. 40) sowie Darnell und Hayhurst (2014). Saavedra (2019) gibt einen Überblick zu Schlüsselkonzepten, benennt und beschreibt ausführlich Studien im Bereich SfE, die postkoloniale Theorien genutzt haben (ebd., S. 211ff.).

²⁹⁹ S. zum Überblick Levermore (2009, S. 41) sowie Darnell und Hayhurst (2014). McDonald (2019) gibt in ihrem Beitrag einen ausführlicheren Überblick zum feministischen Ansatz sowie Einblicke in Studien im SfE-Kontext.

³⁰⁰ S. hierzu zusammenfassend Collison (2019).

³⁰¹ Holt, Deal, Pankow, Pynn und Jorgensen (2019) führen in das theoretische Konstrukt, verschiedene Konzepte und sportspezifische Ansätze ein und stellen ein adaptiertes „model of positive youth development through sport“ vor (ebd., S. 344f.).

Behandlung und die Fokussierung von SfE-Maßnahmen auf marginalisierte junge Zielgruppen aufmerksam, die ‚aktiver‘ als Beratende einbezogen werden sollten. Guest (2019, S. 234) verweist auf die komplizierte Beziehung zwischen Wissens-, Einstellungs- und Verhaltensänderung: „[...] it is not always the case that understanding [health] risks leads to behaving in ways that prevent risk.“ Über die Hälfte der identifizierten Ansätze in Schulenkorfs systematischer Review (et al., 2016) sind der *Psychologie* zuzuordnen. Guest (2019, S. 230) stellt fest: „This SDP paradox of heavily psychological goals, concepts, language, and theory contrasted with relatively little contribution from psychologists begs an obvious question: why are sport psychologists largely missing from the action?“

Forschende aus dem Bereich des *Sportmanagements* veröffentlichten umfangreiche Beiträge zur theoretischen Fundierung im Kontext von Programmdesigns, Eventmanagement, Volunteer-Management, interorganisationalen Partnerschaften, Organisationskapazitäten und Leadership (vgl. Welty Peachey, 2019).³⁰² Hierbei wurde u.a. die Theorie des Sozialen Kapitals häufig genutzt, um die persönliche und kollektive Entwicklungen bei Teilnehmenden, Stakeholdern und Communities zu untersuchen (bspw. Adams, Harris & Lindsey, 2018; Coalter, 2009; Skinner, Zakus & Cowell, 2008; Schulenkorf, 2013; Magee & Jeanes, 2013; Welty Peachey, Bruening, Lyras, Cohen & Cunningham, 2015)³⁰³, aber auch funktionalistische Ansätze zur Erklärung von Volunteer-Motivation (z.B. Welty Peachey, Lyras, Cohen, Bruening & Cunningham, 2014), Ansätze zur Organisationsentwicklung (z.B. Doherty et al., 2014; Svensson & Hambrick, 2019) sowie zu Servant Leadership (Welty Peachey & Burton, 2017) und Sozialem Unternehmertum (Cohen & Welty Peachey, 2015).

³⁰² Die prominente Rolle von Wissenschaftler*innen aus dem Bereich Sportmanagement begründet Schulenkorf (2017) u.a. in der Wichtigkeit, SfE-Programme strategisch zu entwickeln, zu managen und umzusetzen, damit sie ihre Wirkungen entfalten können. Allein zwischen 2000 und 2014 wurden in wissenschaftlichen Sportmanagementzeitschriften 50 SfE-Artikel veröffentlicht (Schulenkorf et al., 2016).

³⁰³ Einen Überblick zu positiven und negativen Konsequenzen, die in Studien bzgl. des Sozialen Kapitals auf verschiedenen Ebenen assoziiert werden, gibt van Eekeren et al. (2013, S. 41f.). Coalter (2007a/b; 2013a) und Skinner, Woolcock und Yoon (2019) geben eine umfangreiche Definition und beleuchten die Rolle des Sports „in generating development and social capital“ (ebd., S. 300).

Die Theorien des *Sozialen Kapitals* und des *Positive Youth Development (PYD)* wurden als jene identifiziert, die im Kontext von SfE-Studien am häufigsten als theoretische Rahmung genutzt werden (Schulenkorf et al., 2016). Darnell et al. (2018a, S. 138f.) bezeichnen diese Zugänge als „relativ neutral“ und „apolitisch“ und dadurch auch „weniger herausfordernd oder ‚bedrohlich‘“ im Hinblick auf den Status Quo der EZ-Landschaft. Doch „connecting sport to development IS political“ (Darnell & Dao, 2017, S. 3). Demnach wird für eine *deskriptivere, kritischere und politischere Auseinandersetzung* i.R. der SfE-Forschung plädiert, indem *politisierende Ansätze* (bspw. postkoloniale oder feministische), „that draw attention to the roots of inequality“ (ebd., S. 139) genutzt werden, wobei allerdings ‚das theoretische oder methodische Rad‘ nicht neu erfunden werden muss“ (Darnell, 2012, S. 5).

Da im SfE-Kontext eine umfangreiche theoretische Basis verschiedener Disziplinen vorhanden ist, sprechen sich Forschende für das *Zusammenführen unterschiedlicher theoretischer Modelle* aus (vgl. Burnett, 2015, S. 388). Einschlägige Zugänge und Modelle sind bereits Ausgangspunkt für die theoretische Weiterentwicklung im Forschungsfeld SfE, bspw. für die *Sport for Development Theorie (SFDT)* (Lyras & Welty Peachey, 2011, 2018) oder dem *S4D Framework* (Schulenkorf, 2012) – mit dem Ziel, die *black box*, d.h. Prozesse und Mechanismen zu erklären, durch die SfE-Programme ihre Wirkungen entfalten (Coalter, 2013a). Beide Konzepte wurden von Welty Peachey et al. (2019, S. 3) als „key theoretical and conceptual developments“ in der bestehenden SfE-Literatur identifiziert. Die Konzepte Sugdens (2010), Schulenkorfs (2012) sowie Lyras und Welty Peacheys (2011) betonen die Wichtigkeit, Zusammenhänge zwischen Programmwirkungen (Mikro-Ebene), Effekten und möglichen Einschränkungen auf Meso- und Makro-Ebene nachzuvollziehen (vgl. Giulianotti, 2019, S. 427). Die wichtigsten theoretischen Schlüsselkonzepte sollen an dieser Stelle in Kürze erläutert werden:³⁰⁴

Das *Ripple-Effect Modell* von Sugden (2010, 2014) stützt sich auf Ansätze des kritischen Linksrealismus und Bottom-up-Aktivismus. Das ‚Welleneffekt‘-Modell veranschaulicht,

³⁰⁴ Welty Peachey et al. (2019) führen fünf theoretische Schlüsselkonzepte auf. Da das *Sport-for-Health Model* (Schulenkorf & Siefken, 2019) u.a. auf zwei der anderen Konzepte basiert (*SFDT* und *S4D Framework*), themenspezifisch ist und von den Autor*innen eher als Ergänzung zu anderen theoretischen Modellen gesehen wird, wird es an dieser Stelle nicht näher beschrieben.

dass Wirkungen einer Maßnahme weniger werden und schwieriger nachzuvollziehen sind, je weiter sie sich von dieser entfernen (bspw. im Familien- oder Community-Kontext). Das Modell wurde von Wissenschaftler*innen bereits häufig genutzt, um den sozialen Hintergrund und Kontext von Programmen und Projekten sowie soziale Outcomes auf unterschiedlichen Ebenen (mikro, meso, makro) zu erklären (bspw. Burnett, 2013a, 2014; Sherry, Schulenkorf, Seal, Nicholson & Hoye, 2017a).

Die *Sport-for-development-theory (SFDT)* nach Lyras und Welty-Peachey (2011) basiert u.a. auf Konzepten der Organisationstheorie und gruppeninterner Kontakttheorie und besagt im Kern, dass „sport initiatives can facilitate personal development and social change by embracing non-traditional sport management practices through an interdisciplinary framework, blending sport with cultural enrichment“ (ebd., S. 313). Die SFDT beinhaltet fünf Bausteine: (1) Wirkungsmessung, (2) Organisation (3) Sport- und Bewegungsaktivitäten, (4) Bildung und (5) kulturelle Bereicherung. Die Theorie wurde bislang mehr erwähnt als genutzt (bspw. von Hancock, Lyras & Ha, 2013; Marshall & Barry, 2015) und häufig weiterentwickelt (z.B. Schulenkorf & Siefken, 2019).

Das *Sport-for-Development (S4D) Framework* von Schulenkorf (2012) ist ein prozessorientiertes Management-Tool, das die Planung, das Management und die Hebelwirkungen von SfE-Projekten und Events unterstützen und erklären soll. Bislang wurde das flexible Konzept zwar oft erwähnt, allerdings in lediglich drei Studien angewendet (Reis, Vieira & Sousa-Mast, 2016; Schulenkorf & Siefken, 2019; Svensson & Levine, 2017).

Coalter (2013b) nutzt zur Strukturierung seiner *programme theory* das *Logische Modell* und stützt sich zusätzlich auf weitere Modelle, die erklären, wie Programme funktionieren und ihre Wirkungen entfalten können (bspw. Faktoren zur Programmgestaltung, die die Resilienz von Jugendlichen fördern oder die wichtige Rolle von Mentoring). Es soll SfE-Praktiker*innen dazu befähigen, sich strategisch damit auseinanderzusetzen, wie sie ihre Programme personell besetzen, konzipieren und implementieren sollen, und wie sie Wirkungen definieren und erreichen können (vgl. ebd., S. 610). Auch diese Theorie wurde bislang wenig angewendet (bspw. Hills, Velásquez & Walker, 2018), was Welty Peachey et al. (2019, S. 8) wie folgt begründen und die Nutzung propagieren: „Overall, and perhaps due to its more recent publication, scholars have not yet engaged fully with

programme theory as a framework. As such, ample opportunity exists to apply, develop, extend, or challenge it in myriad SFD contexts.”

Zudem sollen das *Model for sport for development* nach van Eekeren et al. (2013, S. 38ff.), das mittels einer „relation chain“ den Beitrag von Sport zu übergeordneten Zielen veranschaulicht,³⁰⁵ sowie das i.R. der umfangreichen *Designed to Move*-Studie entwickelte *Human Capital Model* (Nike, 2012), welches auf Literatur zu *Physical Health und Well-being, Individual und Social Capital* beruht (ebd., S. 111, 121), als weitere theoretische Modelle, die im SfE-Kontext entwickelt wurden, an dieser Stelle Erwähnung finden.

Die multidisziplinäre Betrachtungs- und Herangehensweise i.R. der theoretischen Fundierung führt eine vielfältige Nutzung an (theoriegeleiteten und partizipativen) Ansätzen sowie Forschungsdesigns und -methoden mit sich, auf die das Folgekapitel eingeht.

5.4.3 Genutzte Ansätze im SfE-Kontext

Anders als bei der Klassifizierung von Forschungsarbeiten nach Gadais (2019), der Evaluation lediglich in Verbindung von Managementprozessen sieht, werden in Anlehnung an die theoretisch-konzeptionelle Rahmung der Arbeit die unterschiedlichen Ansätze im SfE-Kontext aus der Perspektive der Evaluationsforschung herausgearbeitet.

Nachdem anfänglich hauptsächlich Fallstudien auf Mikroebene (d.h. auf Projektebene, also *Evaluationen von einzelnen Maßnahmen oder Teilbereichen von Maßnahmen*) angesetzt waren (vgl. Cronin, 2011, S. 35) und erste weitblickendere Versuche unternommen wurden, „das Feld abzubilden/to map the field“ (bspw. Cronin, 2011; Hillyer, Zahorsky & Munroe, 2011; Levermore & Beacom, 2009; van Eekeren et al. et al., 2013), bezeichnen es Darnell et al. (2018a, S. 134) als „a marker of the field’s maturation“, dass immer mehr Forschende umfassende *systematische (Systematic) Reviews und Metaanalysen* basierend auf bestehender SfE-Literatur durchführen und veröffentlichen (bspw. Holt et al., 2017; Jones et al., 2017; Langer, 2015; Schulenkorf et al., 2016; Svensson &

³⁰⁵ I.R. ihres Modells unterscheiden van Eekeren et al. (2013, S. 45f.) zwischen zwei Ebenen: Bei der ersten Ebene kann SfE Ergebnisse beeinflussen, d.h. wenn notwendige und unterstützende Bedingungen geschaffen werden, werden möglicherweise sog. „intermediate outcomes“ (z.B. die Entwicklung von Life Skills) erreicht. Diese ‚kurzfristigeren Wirkungen‘ können zwar einen Beitrag zu übergeordneten Wirkungen (bspw. Geschlechtergleichheit) leisten, allerdings haben kontextspezifische Faktoren hierbei einen weit größeren Einfluss, worauf SfE-Organisationen oder Projekte/Programme wiederum keinen Einfluss haben.

Woods, 2017; Whitley et al., 2019a; Welty Peachey et al., 2019). Darnell et al. (2019) sind überzeugt, dass sich aus den Ergebnissen aktueller Reviews *wegweisende managementbezogene und politische Handlungsempfehlungen* ableiten lassen, bspw. die Einbettung von Sport „as one element“ in umfassendere Entwicklungsstrategien und in übergeordnete, entwicklungspolitische Strukturen und Prozesse (*Scaling-up*), die Notwendigkeit der Nutzung von Logischen Modellen, die Beachtung von Kontexten jeglicher Form (historisch, geographisch, politisch, kulturell, sozial) und die Durchführung von *qualitativ hochwertigen Wirkungsstudien*: „[...] while more evidence and data will help overall, research that is either poorly reported or explained, and/or vaguely conceptualized, theorized, and analyzed, may in fact have a negative impact on the field and on the overall understanding of the impacts of SfD interventions“ (ebd., S. 12). Trotz der Fortschritte existieren bislang nur *wenige vergleichende Studien* oder solche, die über einen längeren Zeitraum angesetzt werden (*Langzeitstudien*, u.a. auch mit *ex-post* Perspektiven), und bspw. verschiedene SfE-Maßnahmen im Hinblick auf geographische, kulturelle, soziale, politische, entwicklungsspezifische und historische Komponenten vergleichen (Whitley et al., 2019a; vgl. auch Darnell, Whitley & Massey, 2016, S. 575).

Nach Levermore (2011) sind *theoriegeleitete Logical Framework-Ansätze* gemischt mit *partizipativen Ansätzen* in der sportbezogenen EZ weit verbreitet. Die Logischen Modelle sind auch im SfE-Kontext der Kritik ausgesetzt, hauptsächlich output-orientiert, linear und starr sowie von Geberorganisationen vorgegeben zu sein (‘top-down’), mit denen Partnerorganisationen in den Empfängerländern überfordert seien (vgl. ebd., S. 341; Kay, 2012, S. 898; Lindsey & Grattan, 2012, S. 92). Preti (2012, S. 310f.) weist darauf hin, dass bei diesbezüglicher Kritik zwischen dem allgemeinen Ansatz und den tatsächlichen, in Programmen genutzten „Logframe-Matrizen“ unterschieden werden muss, die häufig zahlreiche Defizite aufweisen: Oft bleiben sie „inflexible blueprints“, indem sie vor Projektbeginn festgelegt, einem Projekt ‚übergestülpt‘ und während des Projektverlaufs keine regelmäßigen Anpassungen vorgenommen werden. Zudem liegt der Fokus von Geldgebenden hauptsächlich auf ‚Zahlen, Daten, Fakten‘, d.h. auf Outputs (bspw. Teilnehmerzahlen), die eine quantitative Datengenerierung erfordern. Logische Modelle sind jedoch nicht zwangsläufig unflexible, statische Produkte „required by boards and

funders“.³⁰⁶ Sie können auch flexible, sich stetig im Prozess (weiter-) entwickelnde Rahmungen darstellen „that are participative, collaborative, iterative and developmental“ (Whitley et al., 2020, S. 23).³⁰⁷

Coalter (2006; 2007a; 2009; 2013a; 2013b) plädiert früh für einen „process-led“, „theory-driven“ und „participatory“ M&E-Prozess in der sportbezogenen EZ. Ein sog. prozesssteuernder Ansatz fokussiert sich auf *formativ-gestaltende* Evaluationen mit dem Ziel der Organisationsentwicklung und der Nachhaltigkeit von Maßnahmen. Hierfür wird eine *theoriegeleitete* (*,theory-driven‘*) *Programmlogik* benötigt, d.h. eine ausformulierte und dokumentierte Abfolge von Ursachen und angenommenen Wirkungen, die aufzeigt „where change occurs“ (Coalter, 2013b, S. 594). Von Coalter als *Programmlogik/-theorie* bezeichnet, finden sich in der SfE-Literatur auch die Bezeichnungen *,Theory of Change‘* (vgl. Engelhardt, 2019, S. 130; Elsemann et al., o.J.) oder *,Program Logic Model‘* (vgl. Sherry et al., 2016, S. 164) wieder. Alle formulieren *Inputs, Outputs, Outcomes und Impacts* auf verschiedenen Ebenen und erklären Prozesse und Mechanismen, wie und warum sie sich bedingen (vgl. Coalter, 2013b, S. 608). Weiterentwickelte theoriegeleitete Ansätze zur Steuerung von Maßnahmen wie der ab den 1990er Jahren in der internationalen EZ besonders einflussreiche *RBM-Ansatz* finden erst spät Einzug in die SfE-Literatur: Im Toolkit zur Messung des sportbezogenen Beitrags zu den SDGs (Commonwealth Secretariat, 2019) wird der RBM-Ansatz vorgestellt und Vorzüge erläutert. Nach Whitley et al. (2019b) nutzen nach wie vor wenige SfE-Organisationen Logische Modelle, was die bestehende Kritik in der Literatur (Lyra & Welty Peachey, 2011) und zuvor durchgeführte Reviews bestätigen (vgl. Cronin, 2011; Jones et al., 2017). Wie bereits erwähnt, betont Coalter (2009, S. 71) hierbei die *partizipative Komponente* des Ansatzes, um u.a. „Capacity-building“, „greater ownership“ und „the development of a self-critical and self-improving organizational culture“ zu fördern. Er zitiert in diesem Zusammenhang Weiss (1997), die den Reiz dieses Ansatzes v.a. darin sieht, dass er eine

³⁰⁶ Auch Coalter (2013a, S. 53) wehrt sich gegen den Vorwurf der Linearität von Programmtheorien.

³⁰⁷ Einerseits hat das Bewusstsein dafür, dass die Übertragung von bestimmten Modellen und Methoden in andere Kulturen und Kontexte eine Herausforderung darstellt, im allgemeinen EZ-Kontext zugenommen, andererseits bleibt aber auch die verstärkte Nachfrage nach international anerkannten und vergleichbaren Systemen wie bspw. hinsichtlich der SDGs (vgl. Preti, 2012, S. 313). Das Commonwealth Sekretariat (2019, S. 48f.) betont die *Wichtigkeit lokaler, kultur- und kontextspezifischer Vorgehensweisen*, auch i.R. von M&E über Standardindikatoren.

Konversation zwischen „researchers, practitioners and policy makers“ voraussetzt (Coalter, 2009, S. 72). Whitley et al. (2020, S. 24) befürworten die aktive Unterstützung von beteiligten Organisationen „in setting up their own [...] frameworks [...] or adapt/adopt existing frameworks“, die sowohl der *Legitimierung* gegenüber Geldgebenden nachkommen, sich an nationalen politischen Schwerpunkten ausrichten und organisationale *Entwicklung* und *Lernen* ermöglichen. Kritiker*innen sind der Ansicht, dass selbst die partizipativsten M&E-Systeme negative Einflüsse haben können und dass viele als partizipativ deklarierte Vorgehensweisen nach wie vor „top-down“-Elemente beinhalten, da sie von externen Evaluierenden gesteuert und im weiteren Sinne an den Agenden der Geldgebenden ausgerichtet sind (vgl. Kay, 2012, S. 892, 898; Levermore, 2011, S. 342; Spaaij, Schulenkorf, Jeanes & Oxford, 2018, S. 26).

Dementsprechend fordern zahlreiche Forschende die intensivere *Nutzung von partizipativen, kultur- und kontextsensiblen Ansätzen, Designs und Methoden*, um komplexe Entwicklungsprozesse ganzheitlich verstehen zu können (bspw. Collison & Marchesseault, 2018; Burnett, 2015; Darnell et al., 2016; 2018; Edwards, 2015; Kay, 2009; 2012), auch i.R. von M&E- Prozessen (vgl. Kay, 2012, S. 900; Kidd, 2011, S. 605; Sherry et al., 2017b). Partizipative Ansätze wie die *Participatory Action Research (PAR)*³⁰⁸ (vgl. Burnett, 2008; 2015; Darnell & Hayhurst 2011; Sherry et al., 2017b) werden genutzt und weiterentwickelt, bspw. in Form des *Sport for Development Impact Assessment Tool (S·DIAT)* (Burnett & Hollander, 2007)³⁰⁹, *Participatory Social Interaction Research (PSIR)*³¹⁰ (vgl. Collison & Marchesseault, 2018) und *Post-colonial-feminist Participatory Action Research (PFPAR)* (vgl. Hayhurst et al., 2015). In diesem Zusammenhang wird die

³⁰⁸ „PAR is [...] a particular orientation to inquiry that encapsulates principles of ethical representation, partnership, transparency, and ownership, in order to extend the rights and hear the voices of those involved“ (Sherry et al., 2017b).

³⁰⁹ Der konzeptionelle Rahmen des S·DIAT basiert u.a. auf methodischen und theoretischen Ansätzen der Evaluationsforschung (u.a. Rossi et al., 2004) mit dem Fokus auf Social Impact Assessment, „[...] the anthropology of development (dealing with impacts across time and potentially-affected systems of human environment) and strategic management“. Zudem wurde mit dem Fokus auf „empowerment through research“ ein „Participatory Action Research (PAR) approach“ adaptiert (Burnett & Hollander, 2007, S. 123).

³¹⁰ „PSIR is a form of participatory research that centralizes, prioritizes and situates sporting interventions and development through the experiences and daily lives of participants. [...] PSIR is focused on gaining greater understandings of cultural systems and social complexities through active roles as researchers from within the fieldwork setting“ (Collison & Marchesseault, 2018, S. 229).

Wichtigkeit hervorgehoben, „[...] for the purpose of empowerment, cultural understanding and insight beyond policy, practice and evaluation“ (Collison & Marchesseault, 2018, S. 226), u.a. durch *Empowerment Evaluationen* lokalen Personen, Programmmitarbeitenden und -teilnehmenden ‚eine Stimme zu geben‘ („empowering local voices“) (vgl. Baker, Baker, Atwater & Andrews, 2015) oder mittels *realistischer (realist) Evaluationen* zu verstehen, „what works for whom in what circumstances and why“ (vgl. Coalter, 2013b, S. 596; Harris, 2018, S. 797; Giulianotti et al., 2019, S. 427).³¹¹

Spaaij et al. (2018, S. 26) stellen sich in diesem Zusammenhang die Frage, *was Forschung eigentlich partizipativ mache*. Sie sind der Ansicht, dass die bisherigen Debatten wenig über die Einigung, dass politische, soziale, kulturelle und ökonomische Aspekte bedacht werden müssen hinausgingen und Verhältnisse von Macht und Wissensgenerierung zu wenig Beachtung fänden. Die Autor*innen untersuchen als partizipativ anerkannte Studien aus dem SfE-Kontext mittels eines mehrdimensionalen Modells³¹² und kommen zu dem Ergebnis, dass selbst diese Defizite ausweisen: „[...] our research does not yet fully embody the participatory principles we are advocating“ (ebd., S. 35). Auch wenn partizipative Forschungsansätze gewollt sind, sind sie in der Durchführung häufig durch institutionelle Abhängigkeiten oder Auftraggeber limitiert (ebd.). Besonders die Reflexion und das kritische Hinterfragen der eigenen Forscherrolle (*Reflexivität*) wird bei SfE-Studien nach Ansicht von Darnell et al. (2018b) bislang zu wenig berücksichtigt. Im Rahmen des qualitativen *SPIDS (sport in development settings)* Ansatzes (Schulenkorf, Edwards & Hergesell, 2020) wird der Reflexion und Reflexivität eine besondere Rolle eingeräumt (vgl. ebd., S. 61).

In ihrer Systematic Review von SfE-Studien identifizieren Schulenkorf et al. (2016, S. 35) ein „potpourri of research approaches and methods“, wobei insgesamt *mehr qualitative Designs* genutzt werden und *weniger quantitative oder Mixed Methods-Designs*. Jedoch

³¹¹ Giulianotti et al. (2019, S. 428) sehen in der Nutzung von diesen Ansätzen eine Möglichkeit „knowledge gaps in SDP between micro-level program success and macro-level change“ zu schließen.

³¹² Spaaij et al. (2018) entwickeln ein dreidimensionales Modell, das nicht nur den *Grad an Partizipation* eruiert, d.h. zu welchem Ausmaß Teilnehmende in allen Phasen des Forschungsprozesses aktiv beteiligt sind, sondern auch den *Grad an Machtverschiebung*, der die Kontrolle und Machtverhältnisse während des Forschungsprozesses beschreibt, sowie den *Grad an Reflexivität*, was u.a. die Reflexion der eigenen Forschendenrolle meint, und damit verbundene Voreingenommenheit und Neigungen sowie die daraus resultierenden Auswirkungen auf die Forschung an sich (ebd., S. 27f.).

weisen sowohl quantitativ als auch qualitativ angelegte SfE-Studien zahlreiche Defizite oder Einschränkungen in ihrer Qualität auf, bspw. im allgemeinen Design und in fehlenden methodischen Details (vgl. Coalter, 2009, S. 56; Darnell et al., 2019, S. 8; Whitley et al., 2019a, S. 181), die teilweise auch von den Forschenden selbst erkannt werden (vgl. Schulenkorf et al., 2016, S. 33). Auch wenn sich die Forschenden der praktischen Herausforderungen bewusst sind, wird zunehmend auch vom bestehenden *Mangel und dem Bedarf an rigorosen Studiendesigns* (d.h. *experimentelle und quasi-experimentelle Designs*, die das Kontrafaktische berücksichtigen und eine Generalisierbarkeit zulassen) gesprochen (vgl. Kaufman et al., 2014, S. 176f., 189; Darnell et al., 2019, S. 1, 13, 15; Lindsey & Darby, 2019, S. 799; Massey & Whitley, 2019, S. 181; Whitley et al., 2019a, S. 181; Whitley et al., 2020, S. 21). Hiermit sind sowohl „[...] *rigorous studies* focused on specific SfD interventions that utilise distinct time points, multiple groups, and validated measures“ (Whitley et al., 2019a, S. 191) gemeint als auch „[...] *rigorous academic research* that is humane, micrological, morally guided, embodied and interpersonal calls for an approach that is culturally and contextually specific, and seeks partnership with particular individuals, groups or settings“ (Darnell et al., 2018a, S. 143).³¹³

Befürworter*innen von partizipativen und qualitativen Vorgehensweisen sind der Ansicht, dass diese gut mit dem Forschungsfeld in Einklang gebracht werden können: „[...] this approach to research is likely more challenging and ‚messier‘ than positivist, quantitative methods, this correlates with the understanding that SDP work is inherently complex, complicated, and messy“ (Darnell et al., 2016, S. 573). Für die Datenerhebung kommen bislang vornehmlich ‚traditionelle‘ qualitative Methoden, u.a. Beobachtungen, Interviews und Dokumentenanalysen, zum Einsatz. Das bereits erwähnte „*process-oriented sport in development settings (SPIDS) research framework*“ (Schulenkorf et al., 2020, S. 53) steht für ein aktuelles qualitatives Design, das verschiedene qualitative Methoden kombiniert und hierdurch eine holistische, gründliche Betrachtung ermöglicht. Die Vorteile von *innovativen, kreativen, kulturell angemessenen und technologisch versierten Designs und Methoden* werden gerade erst entdeckt und sind zunehmend ge-

³¹³ Hervorhebungen durch die Autorin.

fragt (vgl. Darnell et al., 2016, 2018a; Schulenkorf et al., 2016; Sherry et al., 2017b). Hiermit sind Methoden gemeint, bei denen innovative Medien oder Technologien (u.a. „visual and arts-based“, bspw. Fotos, Videos, iPads, Social Media), kreative, flexible und vielfältige Befragungsarten und Vorgehensweisen (bspw. „drawings“, „poems“, „stories of change“, „reflective surveys“, „participatory/community mapping“) bei der Datenerhebung und -auswertung zum Einsatz kommen (Darnell et al., 2016, S. 574; Preti, 2012, S. 314; Schulenkorf et al., 2016, S. 35f.; Sherry et al., 2017b, S. 69). Basierend auf Forschungserfahrungen mit qualitativen Methoden in fünf Ländern merken Collison et al. (2016) an, dass eine explorative Vorgehensweise den Forschenden in seiner Selbständigkeit und Fähigkeit einschränkt, Untersuchungen mit dem Hauptziel des Erkenntnisgewinns durchzuführen, wie es Programmevaluationen oder M&E-Prozessen allgemein zugeschrieben wird, und was u.a. auch von Partnerorganisationen erwartet wird (ebd., S. 419): „[...] keeping research ‚pure‘ in this way does run the risk of alienating local actors who likely feel that they gain little from the research encounter“ (ebd., S. 419). Preti (2012) verweist auf die verbreitete Nutzung von *Mixed Methods* in der EZ und ist der Ansicht, dass „whichever tool is used, it should first and foremost serve the purpose of improving the projects - not to please donors but to increase the benefits to target groups“ (ebd., S. 316). Auch Lindsey und Chapman (2017, S. 39f.) sowie Giulianotti et al. (2019, S. 414) plädieren für die Nutzung gemischter Methoden „[...] ranging from randomized controlled trials (RCTs) through participant action research (PAR) techniques“.

5.4.4 Zur Wirkungsdebatte im SfE-Kontext

„This individual level is the level at which most Sport for Development initiatives intervene, and it is the easiest level at which to collect data. As we progress through the levels of analysis, evidence of impact gets messier: we are often assuming a ‚ripple effect‘ from improving individual’s skills or psychological characteristics, rather than being able to demonstrate impact on communities and societies“ (Cronin, 2011, S. 35).

Aufgrund der ‚Sonderstellung‘ im EZ-Kontext als „low-cost, high-impact tool“ (Kay & Dudfield, 2013, S. 5) und daraus resultierenden ‚Fans und Kritiker*innen‘ von SfE ist der Sport wie kaum ein anderes Instrument dem Druck ausgesetzt, zu beweisen, was er bewirken kann und was es für einen Unterschied macht, mit einem sportbezogenen Ansatz zu arbeiten oder nicht (vgl. Levermore, 2011, S. 340). Im Rahmen der Wirkungsdebatte in der sportbezogenen EZ spielen vor allem die „social impacts“ eine besondere Rolle,

die dem Sport als ‚inherent‘ zugeschrieben werden (vgl. Burnett, 2009; Hatton, 2015; Kay, 2009; Spaaij, 2009).³¹⁴

Die Frage nach Wirkungen ist demnach zentral, Wirkungsstudien sind gefragt, sowohl seitens der finanzierenden als auch der implementierenden Organisationen (vgl. Cronin, 2011). Kritisierende beanstanden die Instrumentalisierung von Wirkungsevaluationen seitens der Geldgebenden, die diese hauptsächlich zur *Legitimierung* ihrer Investitionen und als *Beweise* für die Auswirkungen ihrer Programme („proof of effect“) veranlassen und einfacher zu beantwortende und handhabbare Forschungsfragen hierbei kritischen Reflexionen vorgezogen werden (vgl. Burnett, 2015, S. 387). Durchführungsorganisationen haben ebenfalls den Bedarf, lokale Projekte und Programme durch Wirkungsstudien zu legitimieren (vgl. Preti, 2012, S. 310; Schulenkorf et al., 2016, S. 34).³¹⁵

Die Schwierigkeit diese ‚sportgemachten‘ Wirkungen nachzuvollziehen, zu messen oder gar zu isolieren und demnach die Zuordnungslücke zu schließen, wurde schon lange erkannt (Coalter 2013a, 2013b; Cronin, 2011; Levermore, 2011) und z.T. auch für unmöglich erklärt (vgl. Lindsey & Chapman, 2017, S. 38). Sie dennoch sichtbar zu machen, v.a. hinsichtlich der SDGs, wird jedoch zunehmend dringlicher (Commonwealth Secretariat, 2019; Dudfield, 2019; Engelhardt, 2019; Hatto, 2015; van Eekeren et al., 2013). Die Überhöhung der Wirksamkeit von Sport als ein „all-purpose social vaccine“ (Coalter, 2013a, S. 55) regulierte sich im Laufe der Zeit, indem bspw. Wirkungszuschreibungen ‚vorsichtiger‘ formuliert wurden (vgl. Kay & Dudfield, 2013, S. 9; Schulenkorf, 2017, S. 244). Heute wird hauptsächlich von *direkten oder indirekten Beiträgen* zu bestimmten SDGs gesprochen (vgl. Lindsey & Chapman, 2017).³¹⁶ Dennoch wird davor gewarnt, Wirkungen von SfE-Maßnahmen zu überschätzen, v.a. wenn sie kontrafaktisch formuliert

³¹⁴ In einer Sonderausgabe der Zeitschrift *Sport in Society* (2009, 12/9) wird auf die sozialen Wirkungen von Sport i.R. von zehn Beiträgen ausführlich eingegangen.

³¹⁵ Für Darnell et al. (2018a) ist es keine Überraschung, dass sich „researchers, organizers and communities“ auf die Wirkungen konzentrieren, sehen dies jedoch kritisch: „[...] we suggest that this is likely the case because the logic of neoliberalism (which promotes the importance of cost recovery, instils an audit culture, and privileges efficiency) remains hegemonic and therefore common sense“ (ebd., S. 141).

³¹⁶ „A *direct contribution* may be defined and identified where a sport-based approach has an identifiable impact, independent of other factors, by, for example, [...] making a direct contribution to the achievement of equal opportunities for women through specific employment practices. An *indirect contribution* may be achieved alongside other activities beyond the realm of sport. In such circumstances, determining that sport has made a tangible contribution may require the collation of contributory evidence towards a

werden, bspw. i.R. von Programmangeboten: „The risk of over-claiming the impact of interventions can be increased when impact indicators are based on predictions of what ‘might’ have happened if the intervention had not taken place” (ebd., S. 34).

Das *Fehlen von Evaluationen auf einer Meso- und Makroebene* sowie von *Langzeitstudien, inklusive ex-post Evaluationen*, die sich auf summative Wirkungen fokussieren, wird von vielen SfE-Autor*innen anberaumt (vgl. Burnett, 2015, S. 388; Cronin, 2011, S. 39; Collison et al., 2019, S. 2; Schulenkorf et al., 2016, S. 22, 34; Levermore, 2011, S. 340; Lindsey, 2017, S. 802; Whitley et al., 2019a, S. 190; Zanotti & Stephenson, 2019, S. 172). Nach Schulenkorf et al. (2016, S. 30f.) werden die meisten SfE-Studien auf „community-level“ angesetzt (mit dem Fokus auf lokalen SfE-Initiativen), nur 10% beinhalten Analysen auf einem „international level“ (ebd., S. 34). Lindsey (2017, S. 802) erläutert, dass zwar Ansätze auf der Makro-Ebene im SfE-Kontext genutzt werden, bspw. zur Erklärung von Machtverhältnissen und -gefällen, diese sich jedoch empirisch nur auf die Mikro-Ebene, also auf die Projektebene, beziehen. Evaluationen auf der Meso-Ebene, „[...] especially those that draw on empirical data across a range of institutional actors in the field [...]“, seien deutlich rarer (ebd., S. 802). Einige Anhänger*innen qualitativer Forschung lehnen eindeutige Mikro-Makro-Unterscheidungen ab (vgl. Darnell et al., 2016, S. 574) und begrüßen stattdessen die „messy complexities“ (Thorpe, 2016, S. 556) sowohl auf Mikro- als auch auf Makroebene.³¹⁷

Was Coalter (2013a, S. 45) auch als „Politik-Lobbying“ bezeichnet, äußert sich in der Annahme, dass Verhaltensänderungen auf einer individuellen Ebene (*Mikro*) automatisch in andere Lebensbereiche übertragen werden und dass eine Änderung der Einstellung und des Verhaltens von genug Individuen zu einer kumulativen Wirkung auf der gesellschaftlichen Ebene (*Makro*) führen kann (vgl. Coakley, 2011, S. 309; Cronin, 2011, S. 35; Massey et al., 2015, S. 21). So einfach ist es allerdings nicht, es bedarf bspw. programmatischer Bemühungen auch auf der Makro-Ebene: „[...] the cumulative impacts of multiple micro-level, individually or locally-based programs will not create sustainable

stated goal. Examples may include the use of evidence on the contribution of sport-based activities to specific aspects of public health education” (Lindsey & Chapman, 2017, S. 35).

³¹⁷ Thorpe (2016) arbeitet bspw. mit einem vergleichenden thematischen Ansatz im lokalen und globalen Kontext.

change on the macro-level unless programs deliberately link efforts at the local level to action at the societal level" (ebd., S. 22). Hatton (2015, S. 4) erarbeitete hierzu ein Modell „for scaling-up the social impact“. Es erklärt, wie soziale Wirkungen, die auf der Ebene von SfE-Maßnahmen erzielt werden, auf eine gesellschaftliche Ebene übertragen werden können.³¹⁸

³¹⁸ Durch die drei strategischen und ineinandergreifenden Prozesse „*scale-together*“ (beinhaltet u.a. die Koordination und Bündelung von Partnern und deren SfE-Expertise), „*scale out*“ (bezieht sich auf den Reifungsprozess von SfE-Programmen) und „*scale up*“ (meint die strategische Verbreitung von SfE, bspw. auf weitere Begünstigte, Regionen usw.) sollen auf einer Makro-Ebene Wirkungen erzielt werden (Hatton, 2015, S. 10ff.).

5.5 SfE-Evaluation und -Forschung in Deutschland

Auch wenn Sport bereits früh ein Instrument in der Auswärtigen Kultur- und Bildungspolitik und in der Entwicklungspolitik war (vgl. Kap. 2.1), waren und sind deutschsprachige (sport-)wissenschaftliche Beiträge wenig existent. Frühe Publikationen beschäftigen sich hauptsächlich mit den *Legitimationsargumentationen und Zielen* sowie den *Funktionen* der sportbezogenen EZ (vgl. Giebenhain, 1990, S. 189ff.). Nach Giebenhain (1990, S. 196) haben sich die Elemente der Legitimation wie bspw. entwicklungs- und bildungspolitische Argumente „[...] durch die im Sport vermuteten und z.T. beobachtbaren positiven Wirkungen unter Ausschluss der negativen herausgebildet und über Jahre hinweg als Strukturen manifestiert [...]“.

Digel (1989) versucht früh über seine Beiträge, nicht nur zu legitimieren, sondern auf einer Metaebene Legitimationsmuster zu betrachten. Er kritisierte die Übertragung der positiven Wirkungen auf sozialpädagogischer, sozialstruktureller und politischer Ebene, ohne dabei negative Erfahrungen und der damit einhergehende Einsatz des Sports als „Universaltherapie“ in Betracht zu ziehen (ebd., S. 251). Basierend auf anderen Studien identifizieren Digel und Fornoff (1989) später *sechs zentrale Funktionen des Sports* in der EZ: Persönlichkeitsbildung, Integration, Gesundheit, Chancengleichheit und Grundbedürfnisbefriedigung. Neben dieser *Evaluationssynthese* untersuchen sie i.R. ihres Forschungsvorhabens „Sport in der EZ“, das seitens des BMZ in Auftrag gegeben wurde, die aktuelle Situation des Sports in Ländern der „Dritten Welt“, die Struktur und Organisation der sportbezogenen EZ in Deutschland und vergleichend auch in anderen Staaten (*Review*). Sie konstatieren Ende der 1980er für Deutschland Probleme in der Konzeptentwicklung sowie in der Organisation und Steuerung, letzteres zum einen innerhalb des BMZ, zum anderen auch hinsichtlich der unterschiedlichen Akteur*innen (bspw. BMZ - AA) und der Durchführungsorganisationen (GTZ). Im BMZ würde der Sport scheinbar „mehr oder weniger willkürlich“ „von einem Sektorreferat zum anderen hin- und

hergereicht“, „die Zuordnung [...] scheint sich vorrangig an formalen, seltener an fachlichen Kriterien zu orientieren“ (ebd., S. 187f.).³¹⁹ Abgeleitet aus den Erkenntnissen erarbeiten sie abschließend ein „Sektorkonzept Sport“, indem „Schwerpunkte und Prioritäten sowie Wege zur Erreichung der Ziele [...] aufgezeigt werden“ (ebd., S. 353ff.). Inhaltliche Ideen, die i.R. des Sektorkonzepts bereits zu diesem frühen Zeitpunkt erarbeitet wurden, wurden letztendlich erst mit der Gründung des SV SfE 2013 z.T. in die Tat umgesetzt. Ob bei der Ausgestaltung des SV tatsächlich auch Ergebnisse der Studie genutzt wurden, ist fraglich. Die Studie stellt die erste umfassende Forschungsarbeit zur sportbezogenen EZ dar (Jacobsen, 2015). Hinsichtlich des Nachweises von Wirkungen im SfE-Kontexten halten die Autor*innen (ebd., S. 74) fest:

„Es wird auch künftig nicht möglich sein, die Wirkung des Sports im naturwissenschaftlichen Sinne nachzuweisen. Sport taucht immer im gesellschaftlichen Kontext, d.h. im Verbund mit vielfältigen anderen Variablen auf, von denen Sport nicht isoliert werden kann. Somit ist auch eine eindeutige Zuordnung von Ursache und Wirkung [...] in der Situation des Sports in der Dritten Welt nicht gegeben.“

Demnach sehen sie „[...] die Notwendigkeit einer theoriegeleiteten Diskussion über die Funktion des Sports, um zu fundierten Begründungen und Aussagen über die Wahrscheinlichkeit bestimmter Wirkweisen des Sports in Ländern der Dritten Welt zu erlangen“ (ebd., S. 74). Digel und Fornoff (1989) eruieren auch die Grundzüge der sportwissenschaftlichen Diskussion, die, ausgelöst durch Eplers Kritik an der deutschen Sportförderung Anfang der 1970er, auch vermehrt kritische Beiträge, bspw. aus anthropologisch-kulturrelativistischer (Eichberg, 1973) und lerntheoretischer Sicht (Eberspächer & Egger, 1978), entstehen ließ. Auch wenn sich die i.R. des Sportausschusses angehörten Sportwissenschaftler*innen („Sachverständige“) Ende der 1970er weitgehend einig sind, dass Evaluation ein wichtiger Bestandteil der sportbezogenen EZ sein sollte (vgl. Dt. Bundestag, 1978), und sie Bildungs- und Entwicklungsprozesse „[...] mit dem Ziel Entscheidungen treffen zu können“ bewerten soll (Eberspächer & Egger, 1978, S. 18), ist die Sportwissenschaft zu diesem Zeitpunkt aus Sicht Helmut Dignels „überfordert [...], wenn sie im Bereich der Forschung über Sport in Entwicklungsländern aktiv beratend tätig sein

³¹⁹ Digel & Fornoff (1989, S. 307) sprechen in diesem Zusammenhang von einer „Tendenz zur Fragmentierung“ durch die organisatorische Spezialisierung im BMZ und in den Durchführungsorganisationen sowie von „personeller und struktureller Diskontinuität“, die seitens der Partner „als unzuverlässig und inkompetent“ eingestuft werden könnte.

müsste“. Sie beschäftige sich damit nur „fallweise“ oder „zufällig“, was durch die historische Entwicklung der Sportwissenschaften und mangelnder Ressourcen begründet wurde (Dt. Bundestag, 1978, S. 60). Bis dato war *kein* „wissenschaftlich fundiertes Evaluationskonzept³²⁰ für Entwicklungsprojekte, das dem spezifischen Charakter der Sportwissenschaft Rechnung trüge, [...] bekannt“, wenn überhaupt denkbar (Kirsch, 1978, S. 3), was u.a. der Komplexität der sportbezogenen EZ-Maßnahmen zugeschrieben wurde (vgl. ebd., S. 3; Digel, 1978, S. 18f). Die Entwicklung eines Kriterienkatalogs bzw. eines Evaluationsmodells, abgeleitet aus praktischen Projekterfahrungen, scheint den Sachverständigen allerdings möglich und erwünscht (vgl. u.a. Eberspächer & Egger, 1978, S. 20).³²¹ Neben dem Diskurs rund um Evaluationsmöglichkeiten spricht Digel (1978, S. 15) in seiner Stellungnahme „Erwartungshaltungen“ seitens der Auftraggebenden an, die „mit den Möglichkeiten und Maßstäben der Sportwissenschaft nicht verträglich“ seien: „Durch ein wissenschaftliches Votum wird lediglich versucht, die politische Entscheidung symbolisch zu objektivieren.“ Durch die Aussagen wird zum einen die Fokussierung auf die *Legitimationsfunktion* kritisiert und zum anderen die *Beeinflussung* durch die Auftraggebenden angesprochen.

Mit einem Sammelband wollen Andresen, Rieder und Trosien (1989) eine „Bilanz und Gesamtschau“ der „internationalen Sportbeziehungen zur Dritten Welt“ leisten. Neben Erfahrungsberichten, bspw. aus Sicht von Verbänden oder zu einzelnen Ländern, kommen auch Durchführungsorganisationen zu Wort, u.a. die GTZ (John, 1989) und der DFB (Schmidt & Florin, 1989). Neben ‚analytischen‘ Beiträgen bspw. mit interkultureller Lernintention (Digel, 1989) oder aus der Bildungsperspektive (Gall, 1989), wird auch „[...] das Problem der Expertenauswahl und -vorbereitung diskutiert“ (Wangemann, 1989, S. 235).³²² Im Hinblick auf *Evaluationsprozesse* zu diesem Zeitpunkt stellt Kröger (1989, S. 127) fest, dass diese i.R. von Projektdokumentationen *nicht ersichtlich* werden und dass „es offensichtlich bisher versäumt worden [ist], ein diagnostisches Instrumentarium zu

³²⁰ Hervorhebungen durch die Autorin.

³²¹ S. hierzu auch weitere Stellungnahmen in Dt. Bundestag (1978).

³²² Wangemann (1989) formuliert fachliche und außerfachliche Qualifikationen für Expert*innen in der sportbezogenen EZ, wozu er auch entwicklungspolitische Anforderungen zählt, wie bspw. die „Fähigkeit zum Erfassen entwicklungspolitischer Dimensionen“ (ebd., S. 245). Er konkludiert, dass „nicht die großen Konzeptionen und Politiken der Entwicklungszusammenarbeit im Sport“, sondern die Expert*innen über „Erfolg oder Misserfolg von Sportprojekten entscheiden (ebd., S. 249).

entwickeln, welches für die Bewältigung der Evaluation i.R. der Sportförderung tauglich ist.“ Er sieht die Funktion von Evaluation in der *Optimierung* und *Legitimation* von Projekten. Gall (1989) verweist auf die Nutzung von systematischen und zielorientierten Planungsverfahren, wie der ZOPP der GTZ als *theoriegeleitetem Ansatz* in Anlehnung an das *Logische Modell/Logical Framework*.³²³ Andresen et al. (1989, S. 14) fassen zusammen, „[...] dass wir auch heute noch viel zu wenig wissen, was die Förderung von Sport in der Dritten Welt ist oder bedeuten kann und welche Wirkungen von ihr ausgehen können“.

Ab den 1980ern wird zunehmend die Bedeutung traditioneller Bewegungskulturen in die Überlegungen einbezogen (Dietrich, Jost, Schießl & Schmidt, 1986), der damalige Stand der sportwissenschaftlichen Diskussion kritisch reflektiert und entwicklungstheoretisch eingeordnet (Heinemann, 1985).³²⁴ Neben Trosien (1987), der die Organisation der Sportförderung kritisch beleuchtet, wird auch die Dissertation von Giebenhain (1990) bei Digel und Fornoffs Überblick (1989, S. 287) hervorgehoben.³²⁵ Giebenhain (1990) analysiert in einer Art von *Metasynthese* Projekte des AA und BMZ in den 1970er und 1980er Jahren vor dem Hintergrund entwicklungstheoretischer Konzepte hinsichtlich ihres *entwicklungspolitischen Anspruchs*. Er kommt zu dem Schluss: „Die Zusammenarbeit im Sport mit EL [Entwicklungsländern] entspricht generell nicht den Kriterien der entwicklungspolitischen Leitlinien der BRD in Bezug auf die Sportförderung des AA/BMI. Die Maßnahmen des BMZ hingegen erfüllen einige Aspekte dieser Leitlinien“ (ebd., S. 247). Die AA-Maßnahmen schätzt er im Sinne der damalige ‚Entwicklungshilfe‘ als „kontraeffektiv“ ein, die BMZ-Projekte könnten „entwicklungsfördernd wirken [...], wengleich dies im Einzelnen schwer nachweisbar sein dürfte“ (ebd., S. 247).

Mitte der 1990er zählt die Bibliographie *Sport und Sportförderung in Entwicklungsländern* des Bundesinstituts für Sportwissenschaft über 1500 Quellen, „[...] die sowohl über die Situation von Sport, Sportunterricht und Sportwissenschaft in Entwicklungsländern

³²³ S. Ausführungen unter Kap. 4.2.

³²⁴ Es werden nur exemplarische Autor*innen und Werke aufgelistet. S. vollständig Digel und Fornoff (1989, S. 286f.).

³²⁵ Da sie thematisch spezifisch und für diese Arbeit weniger relevant sind, werden Autor*innen und Inhalte an dieser Stelle nicht weiter ausgeführt. S. hierzu Digel und Fornoff (1989, S. 287).

als auch über sportbezogene Kooperationsmaßnahmen [...] Auskunft geben“ (Sonnenchein & Müller, 1996, S. 3). Nach der Recherche in sportwissenschaftlichen Datenbanken stellt Jacobsen (2015) fest, dass sich nach der umfassenderen Auseinandersetzung in den 1970ern und 1980ern später nur noch wenige Publikationen mit dem Thema seiner Dissertation, der sportbezogenen EZ von in Deutschland ansässigen Organisationen, befassen.³²⁶ Groll (2005) widmet sich allgemeiner der internationalen Sportpolitik und versuchte u.a. anhand von Beschlüssen der Vereinten Nationen und der EU ihr Handlungsfeld herauszuarbeiten. Schreiner (2011) betrachtet den Sport neben Kunst und Wissenschaft als zentralen, jedoch zu wenig beachteten Baustein der Außenkulturpolitik. Bauer (2004) nimmt anhand von Breitensportveranstaltungen, die seitens des IOC seit 1985 durchgeführt wurden, eine Standortbestimmung und Einordnung des Breitensports in der EZ vor. Basierend auf den Ergebnissen und entwicklungstheoretischer Überlegungen leitet er allgemeine Handlungsempfehlungen für EZ-Maßnahmen mit einer breitensportlichen Ausrichtung sowie zu einer modifizierten IOC-Förderstrategie ab. Des Weiteren entwickelt Bauer (2004) Evaluationsraster für weitere IOC-Maßnahmen, die u.a. Überlegungen hinsichtlich Diffusionswirkungen und *Nachhaltigkeit im Sinne von Langfristigkeit* beinhalten. Müller (2004) analysiert vor dem Hintergrund von Modernisierungstheorien die Trainerausbildung des internationalen Leichtathletikverbandes in Ländern des Globalen Südens und schlussfolgerte, dass diese zwar eine leistungssportlich akzeptable Ausbildung sei, allerdings keine entwicklungspolitische Wirksamkeit aufweise, was aufgrund der starken sportfördernden Ausrichtung des Programms nicht verwundert. Im „Handbuch Sportpolitik“ (Tokarski & Petry, 2010) wird Sport neben anderen Politikfeldern auch im Kontext der EZ betrachtet. In drei Beiträgen werden die entwicklungspolitischen und humanitären Dimensionen des Sports beleuchtet, u.a. mittels zweier englischsprachiger Artikel zu bereits erläuterten Themen und Ansätzen³²⁷ sowie ein Beitrag zur Rolle des Sports in der Außenpolitik mit wenig Verbindung zu entwicklungspolitischen Inhalten.

³²⁶ In seiner Analyse untersuchte Jacobsen (2015) verschiedene Datenbanken (SPOLIT, SPOFOR, SPOMEDIA), die Datenbank der Zentralbibliothek der Deutschen Sporthochschule sowie sportwissenschaftliche Publikationsreihen (z.B. Zeitschrift Sportwissenschaft).

³²⁷ Coalter zu „limited focus programmes und broad gauge problems“ (s. Fußnote Kap. 5.3.1) und Sugden zum Ripple-Effect Modell (s. Kap. 5.4.2).

Jacobsen (2015, S. 30) merkt an, dass dato „Arbeiten mit empirischen Forschungsansätzen [...] ebenso wenig dokumentiert [sind] wie quantitative oder statistische Ausarbeitungen. [...] Veröffentlichungen zu den kausalen Zusammenhängen, zu Wirkungsanalysen oder gar Evaluationsergebnisse sind kaum zu finden“. Neben der unbefriedigenden Forschungssituation verweist er auch auf die mangelhafte Dokumentation des Engagements deutscher Akteur*innen in der sportbezogenen EZ. Im Rahmen seiner Dissertation analysiert er mittels qualitativer (Inhaltsanalyse) und quantitativer bzw. statistischer Auswertungsverfahren (u.a. Häufigkeits-/Frequenzanalysen) Projektaktivitäten deutscher Akteur*innen (Regierungsorganisationen, Sportorganisationen, Sportwissenschaft, Stiftungen, NGOs) bzgl. ihrer Ziele, inhaltlicher Schwerpunkte, Zielgruppen, Regionen, Ressourcen, Strategien. Aufgrund der Heterogenität der Akteur*innen und Projekte sowie der mangelhaften Datengrundlage sind die Ergebnisse wenig befriedigend oder aufschlussreich. Einzig und allein der Sammelbandbeitrag von Schott und Merkel (2009) greift in den Nullerjahren die Arbeit des AA und BMZ auf, unterscheidet zwischen internationaler Sportförderung (AA) und EZ im Sport (BMZ) und beschreibt bzgl. der MDGs die Funktionen des Sports in der EZ.

An der DSHS stößt SfE bereits früh auf das Interesse Studierender: Erste Abschlussarbeiten werden meist in Projektkontexten verfasst, wo kleinere empirische Fallstudien durchgeführt werden (z.B. Rockenfeller, 2009; Gaerte, 2011), bspw. mit dem Fokus auf dem Einsatz qualitativer Methoden (bspw. Damerius, 2008; Prenzlau & Rordorf, 2008). Durch die wissenschaftliche Begleitung der GIZ-Vorhaben seit 2013 findet das Thema SfE zunehmend auch Einzug in den Lehrbereich. Es entstehen zahlreiche weitere Abschlussarbeiten, auch im Kontext der GIZ-Projekte (z.B. Schmid, 2015; Kroner, 2015). Seit der Einführung des englischsprachigen Masterstudiengangs *International Sport Development and Politics* steigt die Zahl der studentischen Beiträge. Der Sammelband „Sport und internationale Entwicklungszusammenarbeit“ von Petry et al. (2011) fasst zum einen theoretische Beiträge, bspw. wird die dt. EZ im Kontext sportpolitischen Handelns erörtert (Groll & Hillbring, 2011), friedenspädagogischer Ansätze sowie Interkulturalität vor dem Hintergrund sportbezogener EZ-Maßnahmen diskutiert. Zum anderen werden praktische Einblicke in die Arbeit und Sichtweisen seitens internationaler und nationaler sowie staatlicher und nichtstaatlicher Akteur*innen thematisiert, bspw. aus Perspektive

des AA (Wnendt, 2011) und die durch das AA geförderte Auslandstrainerschule Mainz (Steinmann, 2011), aus der Sicht des kirchlichen Hilfswerks MISEREOR (Hippler, 2011) und der Vereinten Nationen (Lemke, 2011). Im Rahmen des YDF-Projekts in Südafrika, das bis dato größte SfE-Projekt der dt. EZ, kooperiert die GTZ/GIZ mit der Universität Johannesburg. Durch die wissenschaftliche Begleitung werden sog. *Stories from the field* festgehalten, die anhand verschiedener Datenquellen zusammengestellt wurden. Mittels eines *Outcome Monitoring*, bestehend aus 46 Indikatoren, und der *qualitativen* und *partizipativen Most Significant Change*-Methode wurden Ergebnisse zu unterschiedlichen Wirkungsbereichen (*result areas*) festgehalten (Burnett, 2012).

Ein weiterer deutschsprachiger Sammelband mit dem Titel „Sport im Kontext von internationaler Zusammenarbeit und Entwicklung“ (Petry, 2020) fasst zahlreiche Themen- und Handlungsfelder (bspw. Migration, sexuelle Gewalt, Menschenrechte, Sportgroßveranstaltungen, Governance) sowie Ansätze und Erfahrungen, auch aus der Sicht von NGOs. Erstmals werden Strategien, Strukturen, Prozesse und Aktivitäten der SfE-Vorhaben der GIZ in Beiträgen skizziert (Klingmann et al., 2020). Schmid und Arnemann (2020) geben einen Einblick in länderübergreifende Umsetzungs- und Monitoringsstandards, die i.R. des RV in Afrika entwickelt und gesetzt wurden: Die „Wirkungsmonitoring-Systeme“ (WoM der GIZ) lassen die Aggregation von Daten und das ‚Aufzählen von Zahlen, Daten, Fakten‘ zu (vgl. ebd., S. 161f.). Im Vergleich zu einem Sammelbandbeitrag von 2009 (Schott & Merkel, 2009) verdeutlicht der DOSB-Beitrag in diesem Band (Grafarend & Gaerte, 2020), dass sich das Verständnis von SfE und die Maßnahmen des Verbandes ausdifferenziert haben. Der Beitrag von Petry, Bauer und Biermann (2020) skizziert in dem Sammelband *methodische Überlegungen zur Wirkungsmessung* von sozialen Sportprojekten im EZ-Kontext (u.a. typische Forschungsdesigns nach Stockmann, 2006) und ermittelt Herausforderungen, bspw. das Zuordnungsproblem.

Außer einer Publikation der Universitäten Bielefeld und Göttingen, die anhand einer *qualitativen Inhaltsanalyse* die Rolle des Sports in der deutschen Entwicklungspolitik untersucht (Schreiner et al., 2020), wurden hierzu keine weiteren Beiträge auf Englisch bzw. international veröffentlicht.

Es kann festgehalten werden, dass die Auseinandersetzung mit SfE-Forschung und -Evaluation in Deutschland etwa Ende der 1980er einen Höhepunkt hatte: Zu keinem späteren Zeitpunkt wurden Funktionen, Ansätze, Theorien und Methoden so umfangreich im Kreis von Sportwissenschaftler*innen diskutiert. Auch wenn die dt. EZ ab etwa Ende der Nullerjahre wieder mehr in sportbezogene Maßnahmen investierte (vgl. Kap. 2.2), wird die Frage, *wie* diese Maßnahmen untersucht werden, kaum beleuchtet. Alleinig der Beitrag von Petry et al. (2020) greift diese allgemein, ohne explizit auf die GIZ-Vorhaben einzugehen, mit dem Schwerpunkt auf der Messung von Wirkungen auf. D.h., dass seit Beginn der umfangreicheren Investition in die sportbezogene EZ sowohl ein *generelles Forschungs- bzw. Publikationsdefizit* bzgl. der deutschen Maßnahmen besteht, als auch ein *Defizit im Hinblick auf die Vorgehensweisen und Schwerpunkte* i.R. von Forschungs- und Evaluationsarbeiten.

6. Evaluation in der Praxis: Kooperation und Begleitforschung im Rahmen des Sektorvorhabens ‚Sport für Entwicklung‘ der GIZ

Von Beginn an legte die GIZ i.R. ihrer sportbezogenen Vorhaben Wert auf die Kooperation mit einer wissenschaftlichen Einrichtung. Das Evaluationsteam der DSHS begleitet das SV SfE seit 2013 über drei Phasen hinweg. Die Aufgaben und Arbeitsbereiche orientieren sich an den jeweiligen Programmzielen des SV, die sich im Laufe der Zeit veränderten (vgl. Kap. 2.2.1). Im Folgenden wird zuerst ein allgemeiner Gesamtüberblick über die evaluationsspezifischen sowie begleitenden Aufgaben und Arbeitsbereiche der wissenschaftlichen Begleitung gegeben. Der Überblick gibt die Struktur der anschließenden Teilkapitel vor, in denen die Aufgaben und Arbeitsbereiche der drei Phasen beschrieben werden, wobei die Evaluationsspezifika im Fokus der Betrachtungen liegt. Das letzte Teilkapitel fasst alle evaluationsspezifischen Tätigkeiten zusammen und betrachtet die Prozesse übergeordnet anhand von Evaluationsstandards.

6.1 Überblick: Aufgaben, Arbeitsbereiche und Beteiligte im Evaluationsprozess

Folgende Abbildung umfasst die Vielfältigkeit an Aufgaben- und Arbeitsbereichen der wissenschaftlichen Begleitung im fortlaufenden Prozess:

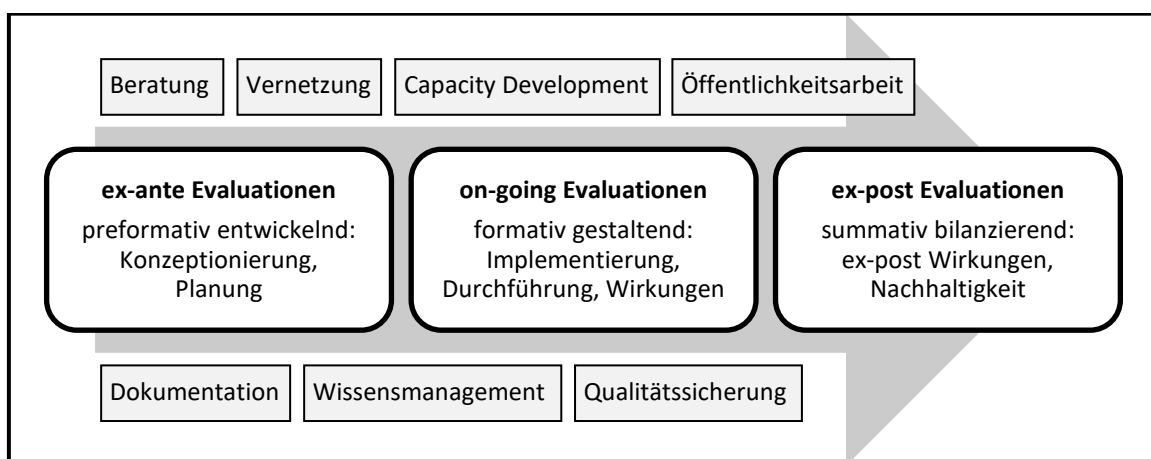


Abb. 11: Aufgaben und Arbeitsbereiche im Prozess der wissenschaftlichen Begleitung (eigene Darstellung nach Schemme, 2003; Stockmann, 2007b)

Mit *evaluationsspezifischen Aufgaben und Arbeitsbereiche* sind all jene Tätigkeiten gemeint, die mit den Evaluationsprozessen im SV SfE in Verbindung stehen, bspw. auch die Entwicklung von Monitoring-Systemen. Im Mittelpunkt steht die Konzipierung und

Durchführung von Evaluationen an sich, die zu unterschiedlichen Zeitpunkten in Bezug auf die Maßnahmen stattfinden (*ex-ante, on-going, ex-post*), die jeweils ein anderes Evaluationskonzept (*preformativ, formativ, summativ*) und ein anderes Erkenntnisinteresse verfolgen, und an denen verschiedene Stakeholder beteiligt sind (vgl. Caspari, 2004, S. 28ff.; Döring & Bortz, 2016, S. 991; Rossi et al., 2004, S. 40; vgl. Stockmann, 2007b, S. 33ff.; 2006, S. 69ff.). An evaluationsspezifischen Prozessen sind/waren nach Brandt (2007) folgende Stakeholdergruppen beteiligt:

- das BMZ als politische*r Entscheidungsträger*in und Geldgebende des SV und der RV SfE,
- die GIZ als Durchführungsorganisation der Vorhaben und zeitgleich Auftraggebende und Geldgebende der wissenschaftlichen Begleitung,
- Programmmanager*innen und -mitarbeitende der GIZ im In- und Ausland,³²⁸
- die wissenschaftlichen Kooperationspartner*innen ICSSPE (Phase 1) und die DSHS als Evaluierende oder Beratende im Evaluationsprozess,
- (NGO-)Mitarbeitende der Partnerorganisationen im Ausland, bspw. als sog. *Gatekeeper* oder aktiv im Evaluationsprozess eingebunden,
- Zielgruppen des Programms (*direkte* Zielgruppe der in SfE geschulten Personen wie Trainer*innen, Lehrkräfte, Sozialarbeiter*innen etc.; Kinder und Jugendliche als *indirekte* Zielgruppe),
- sonstige Betroffene, d.h. Personen oder Gruppen mit wenig oder keinem Einfluss auf die Evaluationsdurchführung, die jedoch mit dieser in Verbindung stehen (bspw. Eltern, die die Einwilligung zu Befragung ihrer Kinder geben müssen).

Neben den Hauptprozessen im Bereich der Evaluation gibt es weitere *begleitende Aufgaben und Arbeitsbereiche*, die z.T. aus den evaluationsspezifischen Aufgaben und Arbeitsbereichen resultieren und/oder diese betreffen und ergänzend dazu verlaufen:

- *Beratung und Vernetzung*: Durch die interdisziplinäre Expertise der wissenschaftlichen Einrichtung(en) und deren weitreichende Vernetzung ist eine Stra-

³²⁸ Für die Partnerländer ist i.d.R. ein*e Ländermanager*innen im Inland verantwortlich, der*die mit einem deutschen Auslandmitarbeitenden und dem lokalen Team im Ausland kommuniziert.

tegie- und Politikberatung sowie die Vermittlung von Expert*innen zu bestimmten Arbeitsbereichen, lokalen Universitäten, Vereins- oder Verbandsstrukturen o.ä. gegeben. Die Beratungsprozesse betreffen verschiedene Adressaten und beinhalten unterschiedliche Schwerpunkte und Themen, zu denen u.a. Recherchen durchgeführt werden. Ein phasenübergreifender Schwerpunkt liegt auf der Unterstützung des SVs bei der Beratung des zuständigen BMZ-Referats.

- *Capacity Development (CD)*: Als Bildungseinrichtung unterstützt die Hochschule beim Auf- und Ausbau von personellen und institutionellen Ressourcen, d.h. sie ist an der Fort- und Weiterbildung von GIZ-Mitarbeitenden und Mitarbeitenden von Partnerorganisationen zu unterschiedlichen Themen beteiligt.
- *Öffentlichkeitsarbeit*: Für die Außendarstellung und im Sinne des Mainstreaming des SfE-Ansatzes werden Ergebnisse und Produkte der Kooperation aufbereitet und genutzt, bspw. i.R. öffentlicher Veranstaltungen wie Podiumsdiskussionen.
- *Dokumentation und Wissensmanagement*: Im Sinne des Wissensmanagements wird SfE-relevantes Wissen (a) mittels Recherchen und Evaluationen erworben (*Erwerb*), (b) basierend auf Ergebnissen und theoretischen Ansätzen aus der Sportwissenschaft innovative Ansätze und Methoden entwickelt (*Entwicklung & Transfer*), (c) Ergebnisse in Form unterschiedlicher Berichtsformate dokumentiert (*Speicherung*) und (d) zur Verfügung gestellt (*Nutzung*) (vgl. Katenkamp, 2001).
- *Qualitätssicherung*: Basierend auf Evaluationsergebnissen werden Handlungsempfehlungen formuliert, die zur Steigerung der Qualität, bspw. von Fortbildungsmaßnahmen, beitragen. Zudem werden Evaluationsergebnisse und theoretische Fundierungen aus dem Bereich der Sportwissenschaften dazu genutzt, standardisierte Leitfäden, bspw. zur Qualitätssicherung von Fortbildungsprozessen zu entwickeln.

In Anlehnung an die unterschiedlichen programmatischen Phasen des SVs SfE (vgl. Kap. 2.2.1) werden im Folgenden drei Phasen der wissenschaftlichen Begleitung, inklusive der Aufgaben, Arbeitsbereiche und beteiligter Gruppen beschrieben. Dabei liegt der Schwerpunkt auf den evaluationsspezifischen Prozessen.

6.2 Phase I: Orientierung

Das SV SfE verfolgte von Beginn an vielfältige Handlungsstränge, um das Thema SfE in der EZ zur Erreichung entwicklungspolitischer Zielsetzungen zu etablieren. Die Kooperation mit wissenschaftlichen Einrichtungen soll die Durchführung und Umsetzung der geplanten Aktivitäten gewährleisten und sicherstellen, auch im Hinblick auf eine Qualitätssicherung und einem neuen Erkenntnisgewinn. Für die wissenschaftlichen Kooperationspartner*innen war die umfangreiche Zusammenarbeit mit der GIZ ‚Neuland‘. Demnach stand die erste Phase (Mitte 2013 bis Mitte 2016) für eine *Orientierung im wissenschaftlichen Begleitprozess*. D.h. zahlreiche Prozesse (bspw. Kommunikations- und Forschungsprozesse) oder Vorgehensweisen (bspw. Priorisierung von Aufgaben, Art und Weise der Berichterstattung) mussten hinsichtlich der Erwartungen der Auftraggebenden erst ‚ausgelotet‘ werden. Inhaltlich lag der Fokus auf der *Konzipierung und Testung strategischer Ansätze* in sieben Pilotierungsländern (Afghanistan, Brasilien, Jordanien, Kolumbien, Namibia, Mosambik, Palästinensische Gebiete). Zudem wurden über die Außenstrukturen und national zahlreiche (Partner-)Netzwerke etabliert, die zum Austausch und zur Positionierung des SV beitrugen.

Während der ersten Phase wurden zahlreiche Evaluationen durchgeführt, v.a. i.R. der pilotierenden Maßnahmen in ausgewählten Ländern. Dabei ging es um die *Planung, Durchführung und Auswertung dieser Maßnahmen*, wozu auch die *länderspezifische Beratung und die Steuerung von Datenerhebungen* gehörten. Eine Evaluation zu bestehenden und abgeschlossenen SfE-Maßnahmen der dt. EZ sowie eine internationale Akteursstudie waren ebenfalls Gegenstand der evaluationsspezifischen Aufgaben.

Internationale und nationale Akteursstudien

Die internationale Akteursanalyse umfasst die Beschreibung von Zuständigkeiten und Schwerpunkten von ausgewählten internationalen Akteur*innen im Feld von SfE und mündet in Empfehlungen zur möglichen Rolle des SV und des BMZ bzgl. zukünftiger Kooperationen oder Beteiligungen sowie hinsichtlich der eigenen Positionierung national und international. Zusammen mit der nationalen Akteursanalyse (vgl. Doll-Tepper et al., 2013), die bereits vor der Gründung des SV durchgeführt wurde, wurden insgesamt 106

nationale und internationale Akteur*innen erfasst (ebd.; Doll-Tepper, Talbot, Petry, Dumon, Weinberg & Koenen, 2014). Da durch die Analysen vorab und mit Beginn des SV sog. *Schlüsselakteur*innen* im Bereich SfE identifiziert werden sollten, sind sie im Bereich von *ex-ante Evaluationen* einzuordnen. Alle Akteur*innen werden Ende 2013 in eine Access-Datenbank eingefügt.

Mit dem Ziel, Partnerorganisationen zu identifizieren oder Länderkonzepte bzw. Programmdesigns zu präzisieren, wurden ebenfalls im Bereich der *ex-ante Evaluationen* länderspezifische Grundlagenuntersuchungen zu bestimmten Akteur*innen und Sportstrukturen erstellt. In Kolumbien wurden 18 Akteur*innen befragt, die im Bereich des Sports und/oder der sozialen (Jugend-)Arbeit tätig sind, darunter NGOs, staatliche Institutionen und Expert*innen (vgl. Petry, Pedemonte, Schröder, Bauer, 2014). Um die Akteur*innen in Schlüssel-, Primär- und Sekundärakteur*innen einordnen³²⁹ sowie ihre Beziehungen zueinander und Abhängigkeiten voneinander veranschaulichen zu können, wurden bei der Ergebnisdarstellung sog. *Akteurslandkarten* genutzt. Zudem stellen *SWOT-Analysen* Stärken (*strengths*), Schwächen (*weaknesses*), Potenziale (*opportunities*) und Risiken (*risks*) bzgl. der wichtigsten Akteur*innen dar. Beide Tools sind Bestandteil des *Capacity Works Handbuchs* der GIZ (vgl. GIZ, 2015a, S. 108, 131). Neben der kolumbianischen Akteursstudie wurde in Mosambik eine Untersuchung zur Sport-, Vereins- und Verbandslandschaft und deren Strukturen, Programme und Verantwortlichkeiten initiiert, bei der 19 Akteur*innen unter Berücksichtigung verschiedener Gesichtspunkte ausgewählt und interviewt wurden (vgl. Merker & Petry, 2014).³³⁰

SfE-Fortbildungsmaßnahmen

Die meisten Evaluationen fanden in der ersten Phase i.R. der *prozessorientierten Begleitforschung* statt und sind damit *on-going Evaluationen*. Insgesamt wurden 23 Studien initiiert, die meisten im Zuge von Fortbildungsworkshops für Coaches. Zum Einsatz kamen hauptsächlich Prä- und Postfragebögen, die die Teilnehmenden zu Beginn und am

³²⁹ Die Kategorisierung in Schlüssel-, bzw. Primäre und Sekundäre Akteure beruht auf Aussagen der Interviews und der allgemeinen Einschätzung der Interviewenden, die sich an Aspekten wie Methodenkompetenz, struktureller Aufbau, Transparenz u.a. orientierten.

³³⁰ Dabei waren beispielsweise die organisatorische und inhaltliche Bandbreite, die Erreichbarkeit und die Abdeckung verschiedener Sportarten von Bedeutung.

Ende der Workshops ausfüllten. Auch Gruppendiskussionen wurden in einzelnen Ländern durchgeführt. Durch den engen Austausch mit lokalem Personal sollten die Methoden und Instrumente an Länderspezifika und kulturelle Kontexte angepasst werden. Basierend auf den Untersuchungsergebnissen wurden (Handlungs-)Empfehlungen, bspw. in Bezug auf die weitere Vorgehensweise im jeweiligen Projekt, ausgesprochen. Die Befragungen zielten zudem darauf ab, Aussagen zu länderspezifischen Zielindikatoren treffen zu können, die sich inhaltlich mit dem Wissenserwerb zu SfE und der Befähigung von SfE-Coaches befassen. Sie dienen damit der Steuerung und sind *formativ-gestaltend*.

Wirkungsnachweis auf Ebene der ausgebildeten Personen: Implementierung der Ansätze und Methoden durch ausgebildete Personen

Um die Implementierung von Inhalten der SfE-Fortbildungsmaßnahmen nachvollziehen und Aussagen zu Veränderungsprozessen i.R. der Trainingsaktivitäten tätigen zu können, wurden sog. *Follow-up Evaluationen* initiiert. Als *on-going Evaluationen* ließen sie mögliche Rückschlüsse auf die Veränderung der Fortbildungsinhalte zu und waren damit *formativ-gestaltend*. Da sie darüber hinaus erste Veränderungen hinsichtlich der Trainingsgestaltung festhielten, können sie auch als *summativ-bilanzierend* eingestuft werden. Im Rahmen der sog. *Follow-up Besuche* wurde ein kombiniertes Untersuchungsinstrument eingesetzt, das aus einem Beobachtungsbogen und einem Interviewleitfaden besteht. Der Beobachtungsbogen beinhaltete Beobachtungsaspekte zur Organisation und zu Inhalten des Trainings, zum Verhalten des Coaches und zur Beteiligung der Kinder und Jugendlichen. Im Interview wurden die Coaches unter Einbezug der Beobachtungskriterien schwerpunktmäßig zur Implementierung der Fortbildungsinhalte befragt, d.h. welche Aspekte bei der Umsetzung schwer bzw. leichter fallen. Das kombinierte Instrument wurde im weiteren Verlauf der Phase adaptiert und in Mosambik, Afghanistan und den Palästinensischen Gebieten eingesetzt.

Wirkungsnachweis auf der Ebene von Kindern und Jugendlichen

Bereits Mitte 2015 wurde seitens der Auftraggebenden (GIZ/BMZ) die Frage gestellt, welche *Wirkungen bei Kindern und Jugendlichen* durch die SfE-Maßnahmen festgestellt werden können. Daher wurden in ausgewählten Partnerländern in Zusammenarbeit mit lokalen NGOs oder internationalen Expert*innen quantitative Befragungsinstrumente

entwickelt. Die Instrumente erfragten neben soziodemografischen Daten auch Wissen, Einstellung und Verhalten hinsichtlich SfE-spezifischer Themen (bspw. Gesundheit, Berufsbefähigung) und wurden zu unterschiedlichen Zeitpunkten eingesetzt (Vorher- und Nachher-Befragung). Neben den quantitativen Untersuchungen wurden erste Erfahrungen mit *qualitativen Erhebungsmethoden* gesammelt: Z.B. wurden im Projekt in Namibia Teilnehmerinnen mittels Tiefeninterviews befragt (vgl. Kroner, 2015). Zudem wurden *Fokusgruppendifkussionen* (eng.: *Focus Group Discussions – FGD*) konzipiert und in einer Pilotbefragung mit zehn Mädchen erfolgreich umgesetzt. Hier sollten die Befragten Wissen-, Einstellungs- und Verhaltens-/Transferfragen hinsichtlich gesundheitsspezifischer Themen und zu SfE-Kompetenzen/Life Skills diskutieren und über ihre Erfahrungen im Projekt sowie dadurch ausgelöste Veränderungen berichten. Da die Evaluationen alle nach einer gewissen Zeit jedoch noch während der Projektverläufe in den Ländern stattfanden und Aussagen über mögliche erste Veränderungen festhalten sollten, werden sie als *summativ-bilanzierende on-going Evaluationen* eingestuft.

Evaluation zu bestehenden und abgeschlossenen SfE-Maßnahmen der dt. EZ

Da es in der dt. EZ bereits einige Erfahrungen mit dem Einsatz von Sport als Instrument zur Förderung von persönlicher und gesellschaftlicher Entwicklung gibt, wurde eine *Projekt-/Maßnahmenanalyse bestehender und abgeschlossener Aktivitäten der dt. EZ* durchgeführt. Durch die Analyse von Projektdokumenten und Interviews mit ausgewählten Personen (Mitarbeitende, Mittler*innen und Zielgruppen) wurden Zielsetzungen, Instrumente, Methoden, Dokumentation und Wirkungen der Maßnahmen aufgearbeitet und ausgewertet sowie aus den Erkenntnissen Empfehlungen für zukünftige Maßnahmen formuliert (vgl. Petry et al., 2015). Da die meisten Projekte zum Zeitpunkt der Evaluation bereits abgeschlossen waren und die Erkenntnisse zur Gestaltung des ‚neuen‘ SVs dienen sollten, wird sie als *performative ex-post Evaluation* eingestuft. Lernerfahrungen, die aus der Dokumentenanalyse und aus Interviews mit Projekt- und NGO-Mitarbeitenden gezogen werden konnten, sind u.a., dass

- die frühe Entwicklung und Etablierung eines gut nutzbaren M&E-Systems inkl. der Einbindung quantitativer und qualitativer Methoden i.R. der Datensammlung von

großer Wichtigkeit ist und zur nachhaltigen Qualitätssicherung der Intervention beiträgt;

- verschiedene Bewegungs- und Sportarten (bspw. auch Trendsportarten) bei der Entwicklung von weiteren SfE-Methoden einbezogen werden sollen, um dem jeweiligen kulturellen Kontext angepasst und zielgruppengerecht arbeiten zu können, sowie schwer erreichbare Zielgruppen wie junge Frauen und Mädchen anzusprechen und zu involvieren;
- alle Projektmitarbeitende neben sportartenspezifischen Kompetenzen über ein umfangreiches pädagogisches, d.h. auch methodisch-didaktisches Wissen verfügen müssen, damit SfE-Inhalte und Methoden zielführend eingesetzt werden können,
- Sport verstärkt als Querschnittsthema genutzt werden sollte, d.h. dass mit vielen unterschiedlichen Akteur*innen, bspw. aus dem Bildungs-, Gesundheits- oder Umweltbereich, zusammengearbeitet werden sollte, um SfE-Ansätze in bislang ‚sportfremden‘ Bereichen zu verbreiten und zu verankern.

Begleitende Aufgaben und Arbeitsbereiche

Ein wichtiger Schritt der *Vernetzung* war die Etablierung des BMZ-Thementeam, das sich seit Anfang 2015 regelmäßig trifft. Neben anderen nationalen Akteur*innen (vgl. Kap. 2.2.2) beteiligen sich die wissenschaftlichen Kooperationspartner*innen aktiv.

Im Zuge von *CD-Maßnahmen* wurden 2014 in Namibia und 2015 in Brasilien GIZ-Mitarbeitende, Mitarbeitende von Partnerorganisationen und Coaches zum Thema M&E geschult. Hier ging es darum, praxisnah wissenschaftliche Prozesse und Methoden erfahrbar zu machen und in direkten Überlegungen auf die eigene Arbeit bzw. auf das eigene Projekt zu übertragen. Konkret wurden die Teilnehmenden u.a. in Erhebungs- und Auswertungsmethoden aus dem quantitativen und qualitativen Forschungsbereich qualifiziert, um Untersuchungen selbst planen und umsetzen zu können.

6.3 Phase II: Strukturierung, Standardisierung und Scaling-up

Inhaltlich lag der Fokus der zweiten Phase (Mitte 2016 bis September 2019) auf der Ausarbeitung und Verfeinerung der länderspezifischen SfE-Ansätze und -Methoden bzgl. der SDGs³³¹, der Verbesserung der Fortbildungskonzepte, der nachhaltigen Verankerung der Ansätze und Methoden in den Partnerstrukturen (*Scaling-up*)³³² und dem Wissensmanagement. Neben den bereits gewählten Partnerländern kamen weitere hinzu, manche Projekte wurden abgeschlossen oder in andere regionale Vorhaben überführt. Am Ende der Phase lagen unterschiedlich umfangreiche Daten zu acht Partnerländern vor (Afghanistan, Brasilien, Indonesien, Kolumbien, Namibia, Palästinensische Gebiete, Region West-Balkan, Uganda).³³³

Im Hinblick auf die evaluationsspezifischen Aktivitäten stand diese Phase für die *Strukturierung und Vereinheitlichung von Evaluationsprozessen*. Der Schwerpunkt lag nicht mehr wie in der ersten Phase auf Evaluationen i.R. der Fortbildungsmaßnahmen in den Partnerländern, sondern auf der Unterstützung bei der Durchführung von *M&E-Maßnahmen zu den Scaling-up Prozessen* in den Partnerländern.³³⁴

Scaling-up: Standardisierung und Strategien

Nachdem in der ersten Phase Probleme bzgl. der einheitlichen Berichterstattung zu den Projekten in den Partnerländern erkannt wurden, was auf die unterschiedlichen Monitoringsysteme zurückzuführen war, fanden Anfang der zweiten Phase umfangreiche Beratungsprozesse hinsichtlich der Etablierung von Monitoringsystemen und der Entwicklung von Standardindikatoren statt. Anfang 2017 wurde ein *länderübergreifendes Standardindikatorensatz* zu den *Scaling-up Schlüsselfaktoren* der GIZ entwickelt (vgl. GIZ,

³³¹ Die länderspezifischen SfE-Methoden waren anfänglich sehr ‚allgemeiner Natur‘, d.h. dass über den Sport sog. *Life Skills* vermittelt werden, die keinen spezifischen SDGs zuzuordnen sind. Erst im weiteren Verlauf haben sich Ansätze und Methoden geschärft und die Maßnahmen ihre intendierten Beiträge zu den SDGs klarer formuliert.

³³² Als *Scaling-up* bezeichnet die GIZ (2015a, S. 227) die „Verbreitung und Verankerung von Erfahrungen, Lernprozessen, Wissen und Lösungen mit dem Ziel, strukturbildende Wirkung und Breitenwirkung von innovativen Pilotmaßnahmen und neuen Ansätzen zu fördern“.




³³³ Zur Region West-Balkan gehören die Projektländer Albanien, Bosnien und Herzegowina, Kosovo, Nord-Mazedonien und Serbien. Auch wenn die Region fünf Partnerländer umfasst, wird im weiteren Verlauf aus Gründen der Vereinfachung von Partnerland gesprochen.

³³⁴ Fast alle Informationen in diesem Teilkapitel basieren auf dem Schlussbericht der DSHS (Petry & Bauer, 2019). Wenn weitere separate interne Berichte zu einzelnen Evaluationen vorliegen, werden separate Quellen genannt.

2015, S. 228): (1) Scaling-up, (2) Ownership der Schlüsselakteur*innen, (3) Mehrebenenansatz, (4) Wirkungsnachweis, (5) Standards und Handbücher, (6) Verbreitungsstrukturen und (7) Zeit- und Finanzierungshorizont. Zu den Faktoren wurden unterschiedlich viele (1-4) Indikatoren formuliert, teilweise auch länderspezifisch. Über das Indikatorenset sollten allgemein Informationen gewonnen werden, ob und inwiefern die entwickelten SfE-Ansätze und Methoden in den Partnerstrukturen verankert sind und im Speziellen Angaben zu ausgewählten Indikatoren der übergeordneten Wirkungsmatrix des SV getroffen werden. Die Zielerreichung der Indikatoren sah in den einzelnen Partnerländern zum Ende der zweiten Phase unterschiedlich aus. Dies lag bspw. daran, dass in einigen Ländern entweder (noch) keine Aussagen zu bestimmten Indikatoren getroffen werden oder diese ‚teilweise erreicht‘ waren, da Projekte erst begonnen oder eine erst sehr kurze Laufzeit hatten oder Instrumente nicht in jedem Land zum Einsatz kamen.

In den Partnerländern wurden unterschiedliche *Scaling-up-Strategien* verfolgt (vgl. Ramboell, 2013, S. 47ff; GIZ, 2015a, S. 86; 2015b, S. 6f; 2014, S. 4): Erstens wurden SfE-Inhalte in nationale Bildungsprogramme, Lehrpläne oder Verbandsausbildungen eingebunden und damit auf eine nationale Ebene übertragen. Durch dieses *vertikale Scaling-up* konnten sich die innovativen SfE-Ansätze und -Methoden flächendeckend im jeweiligen Land ausweiten. Zweitens wurden die SfE-Inhalte auf eine neue Einheit auf der gleichen Ebene übertragen. D.h. dadurch, dass bspw. NGOs die Ansätze und Methoden übernommen haben und diese über ihr eigenes Netzwerk in unterschiedliche Regionen verbreiteten, hat ein *horizontales Scaling-up* stattgefunden. Drittens wurden die SfE-Ansätze und Methoden auch auf andere Kontexte und/oder Sektoren (z.B. Gesundheit, berufliche Bildung) übertragen, was einer *funktionalen Scaling-up-Strategie* entspricht. Tabelle 6 veranschaulicht, dass in den meisten Ländern ein bis zwei, in Uganda sogar drei Scaling-up Strategien genutzt wurden:

Tab. 6: *Scaling-up Strategien in den Partnerländern des SV SfE (eigene Darstellung)*³³⁵

Scaling-up Strategie		AF	BR	ID	KO	NA	PG	RWB	UG
Vertikales Scaling-up		X		X	X	X			X
Horizontales Scaling-up			X		X		X	X	X
Funktionales Scaling-up						X		X	X

Im Folgenden sollen die unterschiedlichen Evaluationen zu den Scaling-up Prozessen erläutert werden. Dabei ist anzumerken, dass die Informationen, die über ein Erhebungsinstrument eingeholt wurden, Aussagen zu unterschiedlichen Standardindikatoren zuließen.

Mehrperspektivische Betrachtung von Scaling-up-Prozessen

Zur Betrachtung der Scaling-up Prozesse aus Sicht der Auslandsmitarbeitenden der GIZ kam ein *Fragenkatalog/Länderbericht* zum Einsatz, der zur Beantwortung offener Fragen einen Lückentext beinhaltete und der in fast allen Ländern jährlich genutzt wurde. Zur Erfassung der Prozesse aus der Perspektive der Kooperationspartner*innen wurde ein Partnerfragebogen genutzt, der offene und geschlossene Fragen enthielt und von insg. 14 Partnerorganisationen aus sieben Ländern ausgefüllt wurde. Da die Evaluationen jährlich und mit dem Ziel stattfanden, den Status Quo der Scaling-up-Prozesse zu überprüfen oder abschließend zu bewerten, werden sie als *formativ-summative ongoing Evaluationen* eingestuft. Über die beiden Instrumente konnte vor allem der zentrale Scaling-up-Faktor *Ownership von Schlüsselakteur*innen* mehrperspektivisch beleuchtet werden. Ownership sollte dann gegeben sein, wenn die länderspezifischen SfE-Ansätze und Methoden zum einen gemeinsam mit bestimmten Partner*innen entwickelt werden und zum anderen fest in die Struktur der Partner*innen integriert sind. Es zeigte sich, dass Ansätze und Methoden in allen Ländern gemeinsam mit Partner*innen entwickelt worden sind. Dabei erfolgte die Einbindung auf ganz unterschiedliche Weise (bspw. über intensive Workshops mit lokalen Expert*innen und/oder die Beratung

³³⁵ Die Abkürzungen stehen für Afghanistan (AF), Brasilien (BR), Indonesien (ID), Kolumbien (KO), Namibia (NA), Palästinensische Gebiete (PG), Region West-Balkan (RWB) und Uganda (UG).

durch externe Expert*innen) und zu unterschiedlichen Zeitpunkten (z.B. bei der Übertragung von Methoden auf andere Sportarten, nachdem sportartenspezifische Manuale durch Verbände entwickelt worden waren). In diesem Zusammenhang kann festgehalten werden, dass durch *partizipative Entwicklungsprozesse* Ownership bei den Partner*innen erzeugt wurde und die Wahrscheinlichkeit der Verankerung der SfE-Methode zunahm. Solche partizipativen Prozesse benötigen *Zeit*, die jedoch durch strukturierte Formate (bspw. Test-/Entwicklungsworkshops) *effektiv gesteuert* werden können. In fast allen Ländern sind Ansatz und Methode fest in die Partnerstruktur integriert, zu zwei Ländern konnten keine Aussagen getroffen werden, da keine Daten vorlagen oder der Prozess im Partnerland noch am Anfang stand. Hier kann festgehalten werden, dass die Verankerung von SfE-Methoden in bestehende Lehrpläne von Schulen oder Universitäten als besonders wirksam und nachhaltig erachtet wird.

Darüber hinaus wurden über die Instrumente weitere Informationen zu den Scaling-up-Faktoren gesammelt, bspw. zu anvisierten Absichtserklärungen und Vereinbarungen hinsichtlich der *Weiterentwicklung, -führung und -finanzierung der SfE-Ansätze und Methoden* in den Ländern nach dem Förderende der GIZ. Eine sehr umfangreiche Beratungsleistung der GIZ fand hierzu bspw. in Brasilien statt: Hier war die finanzielle Nachhaltigkeit des Projektes von Beginn an Thema beim Austausch mit den Partnerorganisationen. Durch Kooperationen mit lokalen Akteur*innen wie Gemeinden, Universitäten, Schulen oder Sportvereinen wurden in fünf Regionen unterschiedliche Strukturen aufgebaut, welche Einsatz und Nutzung des SfE-Ansatzes sicherstellten und durch die die Organisationen neue Einnahmequellen erschließen konnten. Im Sinne eines *Mehrebenenansatzes* haben die SfE-Maßnahmen nach Angaben der GIZ-Mitarbeitenden und einzelner Partner*innen in allen Ländern eine Verknüpfung von Akteur*innen auf verschiedenen Ebenen (national, regional, lokal) gefördert. Auf nationaler Ebene entstanden über die SfE-Projekte bspw. neue Kooperationen zwischen staatlichen Akteur*innen und/oder nationalen Verbänden oder neue Verknüpfungen von unterschiedlichen Akteur*innen durch landesweite Netzwerke. In Uganda wurde von Beginn an ein Netzwerkansatz verfolgt, in dem ein aus verschiedenen nationalen, regionalen und lokalen Akteur*innen bestehendes *Training and Resource Network* etabliert wurde. In der Re-

gion Westbalkan erfolgte durch die Einbindung von Akteur*innen aus fünf Ländern sogar ein transnationaler Austausch: Vertreter*innen nationaler Partnernetzwerke treffen sich inzwischen regelmäßig zum Erfahrungsaustausch, zur Weiterentwicklung ihrer lokalen Strategien, des SfE-Ansatzes und Methode.

Zum Faktor *Verbreitungsstrukturen, Kommunikation und Netzwerkbildung* konnte nach Aussage der befragten Mitarbeitenden der GIZ und von Partnerorganisationen festgehalten werden, dass durch die SfE-Maßnahmen in sieben Ländern ein Netzwerk oder eine überregionale Institution in Bezug auf Kommunikation, Vernetzung und Erfahrungsaustausch gestärkt wurde. Durch die Durchführung von jährlichen internationalen SfE-Lehrgängen von Instructors in Deutschland seit 2016 wurde zudem die Kommunikation und der Erfahrungsaustausch zwischen Personen aus allen Partnerländern gewährleistet. Zudem gaben die Befragten an, dass in fast allen Partnerländern die *entwickelten Referenz- und didaktischen Materialien*³³⁶ seitens der Partnerorganisationen genutzt werden.

Wirkungsnachweis auf Ebene der ausgebildeten Personen: Implementierung der Ansätze und Methoden durch ausgebildete Personen

Um einen Eindruck zu bekommen, welche Auswirkungen die Teilnahme an den SfE-Weiterbildungen auf die Trainings- oder Unterrichtsgestaltung hatte, wurden unterschiedliche methodische Zugänge gewählt: Die bereits in der ersten Phase initiierten *Follow-up-Besuche (Beobachtung und Interview)* wurden ausgebaut, länderspezifisch modifiziert und übersetzt. Darüber hinaus wurden sie im Sinne eines Mixed-Methods-Ansatzes um eine *quantitative Online-Befragung* ergänzt. Beide *on-going Evaluationen* sind *formativ*, da die Ergebnisse Informationen zur weiteren Programmsteuerung liefern und zeitlich *summativ* sind, da nach einer gewissen Laufzeit Rückschlüsse zum relevanten Beitrag der Maßnahmen gezogen werden können, d.h. ob und inwiefern ausgebildete Personen ihre Trainingsinhalte SfE-spezifisch verändert haben.

³³⁶ Mit *Referenzmaterialien* sind jegliche entwickelten Standarddokumente gemeint (bspw. Leitfäden, Anleitungen etc.). Mit *didaktischen Materialien* sind jegliche entwickelten Handbücher, Manuale oder Lehrvideos gemeint, die die länderspezifische Methode (inkl. Übungen und Spiele) beinhalten.

Insgesamt wurden 132 Personen nach ihrer SfE-Weiterbildung in ihrem Trainings- oder Unterrichtsetting besucht. Bspw. wurden in *Afghanistan* die Besuche an eine Verteilung von Sportmaterialien geknüpft, wodurch Lehrer*innen an 48 Schulen in Kabul und Mazar-e-Sharif durch Instructors des Bildungsministeriums supervisiert und bewertet werden konnten. In *Kolumbien* wurden durch eine beauftragte Partnerorganisation insgesamt 40 Personen z.T. zweimalig besucht. In den *Palästinensischen Gebieten* war geografisch die Möglichkeit gegeben, die Besuche zu einem Bestandteil der Fortbildungsstruktur zu etablieren, d.h. über die Hälfte der ausgebildeten Coaches wurden monatlich von der GIZ oder einer Partner-NGO aufgesucht und bei der Umsetzung der Ausbildungsinhalte unterstützt. Dass generell auch in der zweiten Phase insgesamt sehr wenige Follow-up-Besuche durchgeführt wurden, hatte finanzielle, personelle und logistische Gründe, bspw. standen häufig keine zusätzlichen Projektmittel zur Verfügung. Das Untersuchungsdesign, das darauf abzielte, Personen zu zwei Zeitpunkten zu besuchen, war für einige Länderkontexte zu komplex angelegt. Zahlreiche äußere Umstände (große Entfernungen, Wechsel und Fluktuation von Mitarbeitenden, schwierige Sicherheitslage) führten dazu, dass Daten fehlten und Ergebnisse oft schwer vergleichbar waren.

Die quantitative *Online-Befragung* fand 2018 und 2019 statt. Insgesamt waren zu dem Zeitpunkt 3491 Personen in einem SfE-Workshop der GIZ fortgebildet, wovon 788 aus fünf Ländern, d.h. 22,6% aller ausgebildeten Personen, an der Umfrage teilnahmen. Etwa 30% der befragten Personen waren weiblich (Petry & Neeten, Dezember 2019). Bei der Befragung bestand die Möglichkeit, dass hauptsächlich jene Personen teilnahmen, die sich nach wie vor im SfE-Kontext engagieren oder Interesse daran haben. Über einen Großteil der Personen (fast 80%), die ebenfalls an SfE-Fortbildungen teilnahmen, allerdings nicht an der Befragung, kann keine Aussage getroffen werden. Möglicherweise haben sie keinen Bezug mehr zur Thematik und implementieren demnach auch keine Inhalte der Workshops. Vor diesem Hintergrund ist die Stichprobe *nicht repräsentativ* für die Grundgesamtheit der ausgebildeten Personen. Für die Umfrage wurde ein großer Aufwand betrieben, bspw. wurde der Fragebogen in sieben Sprachen übersetzt, in das Programm *Unipark* eingespeist und inklusive Übersetzung ausgewertet, was umfangreiche zeitliche und personelle Ressourcen beanspruchte.

Wirkungsnachweis auf der Ebene von Kindern und Jugendlichen

In Anlehnung an die vorangegangenen Wirkungsevaluationen wurde in Kooperation mit verschiedenen NGOs in Brasilien Wirkungsmessungen über Prä- und Postfragebögen vorgenommen. Bereits bestehende Instrumente wurden i.R. eines CD-Workshops in einem partizipativen Prozess mit den NGOs weiterentwickelt und angepasst. Einstellungs- oder Verhaltensänderungen konnten bei den Kindern und Jugendlichen aufgrund verschiedener Herausforderungen rund um den Datenerhebungsprozess allerdings nicht nachgewiesen werden (bspw. fehlende Kontrollgruppen, Fluktuation).

Standardisiertes Monitoring

Über ein *standardisiertes Monitoring-Template* konnten detaillierte Informationen zu den Fortbildungsmaßnahmen und weiterer Aktivitäten, bspw. zu Veranstaltungen für Kinder und Jugendliche (Summercamps, Turniere etc.), größere Events für/mit Partner*innen (Netzwerktreffen, Kongresse, Podiumsdiskussionen etc.) sowie öffentliche Veranstaltungen und Festakte (Messen, Festivals etc.), gesammelt und aggregiert werden. Im Vergleich zur ersten Phase konnten Ende der zweiten Phase vergleichende und einheitlichere Zahlen generiert und aggregiert werden. Dementsprechend waren in den Partnerländern ca. *3400 Coaches und 280 Instructors* i.R. von 300 Fortbildungsworkshops in SfE-Ansätzen und Methoden geschult worden.

Begleitende Aufgaben und Arbeitsbereiche

Im Rahmen von *Beratungsprozessen* wurden Ende der zweiten Phase alle bislang gesammelten Erfahrungen zu den Wirkungsevaluationen in einem *Grundsatzpapier* festgehalten, das sowohl als Diskussionsgrundlage für das BMZ als auch als Grundlage für weitere Wirkungsevaluationen dienen sollte und mit der *Stabsstelle Evaluierung der GIZ* abge-

stimmt wurde. Das Dokument skizziert Grundlagen zur Wirkungsmessung und gibt anhand der gesammelten Erfahrungen und abgeleiteten Herausforderungen³³⁷ einen Ausblick, wie Wirkungsevaluationen künftig durchgeführt und dabei der Beitrag zu den SDGs fundierter dargestellt werden könnte. Wichtige Erkenntnisse sind, dass sich qualitative Evaluationsansätze als gewinnbringend, aber ressourcenintensiv und der Einsatz quantitativer Instrumente (Prä- und Postfragebögen) methodisch und organisatorisch als herausfordernd herausstellten.

Hauptaugenmerk der zweiten Phase bzgl. der begleitenden Aufgaben und Arbeitsbereiche lag auf der *Dokumentation* und dem *Wissensmanagement*: Bereits bei der Evaluation zu bestehenden und abgeschlossenen Maßnahmen der GIZ (s. Kap. 6.2) wurde deutlich, dass umfangreiches Wissen und Erfahrungen in der sportbezogenen EZ besteht. Durch die Pilotierung in weiteren Partnerländern seit 2013 wurden weiteres Wissen und Erfahrungswerte generiert.³³⁸ Im Zuge der Entwicklung der Homepage *S4D Resource Toolkit* wurden Wissen und Lernerfahrungen aus der Praxis aufbereitet, gebündelt und im Sinne des *Mainstreaming des SfE-Ansatzes* anderen Nutzer*innen und Praktiker*innen zur Verfügung gestellt. Zahlreiche Dokumente aus den Ländern wurden aufbereitet oder neu erarbeitet und stehen jetzt als *Standarddokumente, Leitfäden und Arbeitshilfen* zur Verfügung.³³⁹ Das Toolkit ist seit Ende 2018 online.

³³⁷ Neben allgemeinen Herausforderungen (bspw. fehlende länderspezifische Wirkungsmodelle und theoriebasierte Wirkungshypothesen; Problematiken i.R. der Managementprozesse in den Projekten, die Wirkungen und dazugehörige Baselines nicht von Anfang an einplanen) gab es vor allem Schwierigkeiten bei der Implementierung rein quantitativer, experimenteller Designs: fehlende Kontrollgruppen, keine vergleichbaren Stichprobengrößen u.a. aufgrund von Fluktuation, unterschiedliche Datenerhebungszeitpunkte (t0 war nicht gegeben, zu wenig Zeit zwischen Erhebungszeitpunkten), ‚keine‘ Ergebnisse (aufgrund fehlender Kontrollgruppen, Probleme der Vergleichbarkeit, zu geringe Stichprobengröße etc.) oder ‚negative Effekte‘ (d.h. keine oder negative Verhaltens- und Einstellungsveränderung), fehlende Triangulation.

³³⁸ Bspw. lagen schon zum Ende der ersten Phase zahlreiche Manuale zu den SfE-Methoden, Fortbildungskonzepte und M&E-Instrumente in unterschiedlichen Sprachen vor.

³³⁹ Beispiele für sog. *Standarddokumente*: Kompetenzraster für Coaches und Kinder/Jugendliche, Leitfäden zur Qualitätssicherung bei SfE-Fortbildungsmaßnahmen, zur kompetenzorientierten Durchführung von SfE-Trainingseinheiten, zur Durchführung von Follow-up-Besuchen, zur Berichtserstattung etc.

6.4 Phase III: Capacity Development und Wirkungsmessung

Die inhaltlichen Schwerpunkte der dritten Phase (Mitte 2020 bis September 2022) liegen für die wissenschaftliche Begleitung auf den Bereichen *Capacity Development, Standardisierung, Vernetzung von Universitäten* und *Wirkungsmessung*. Zur Umsetzung von CD-Maßnahmen in unterschiedlichen Sprachen holte sich die DSHS internationale Unterstützung und arbeitet mit externen Beratenden zusammen. Da die Begleitung noch bis Ende 2022 andauert, werden an dieser Stelle lediglich die relevantesten Aufgaben und Arbeitsbereiche skizziert.

Wirkungsnachweis auf der Ebene von Kindern und Jugendlichen

Im Vergleich zu den vorherigen Phasen ist die DSHS nicht mehr aktiv in M&E-Prozesse in den Partnerländern eingebunden, sondern hat in diesem Zusammenhang vielmehr eine beratende Funktion mit dem Schwerpunkt auf der *Beratung i.R. von Wirkungsevaluationen*. Das Thema *Wirkungsmessung* ist in der dritten Phase prominent platziert, d.h. es ist in der Wirkungsmatrix des SVs verankert und durch Indikatoren operationalisiert. Es sollen fünf Wirkungsstudien (*summative on-going Evaluations*) auf Basis von relevanten SDG-Targets durchgeführt werden. Ein weiterer Indikator setzt sich zum Ziel, dass weitere EZ-Vorhaben die entwickelten Instrumente zur Wirkungsmessung nutzen. In ausgewählten Ländern wurden über die kompetenzbasierten Ansätze und daraus abgeleiteten *Impact Indicator Frameworks*³⁴⁰ Wirkungslogiken erarbeitet und Fragebögen entwickelt. Daten sollen zu unterschiedlichen Zeitpunkten bei Interventionsgruppen gesammelt und mit denen von Kontrollgruppen verglichen werden (*quasi-experimentelle Designs*).

Hochschulmapping

Durch eine Evaluation zu Hochschulen in Deutschland und in Partnerländern (*Hochschulmapping*) sollen konkrete Umsetzungsvorschläge zur Förderung von Interdisziplinarität und Hochschulvernetzung im Bereich SfE getätigt werden und ein anwendungsorientiertes Konzept entstehen. Über einen mehrsprachigen Online-Fragebogen mit offenen und geschlossenen Fragen konnten 24 Vertreter*innen von Hochschulen in 13

³⁴⁰ S. exemplarisch zum Thema *Gesundheit (SDG 3)* GIZ und DSHS (2020).

Ländern befragt werden (vgl. Moustakas, Petry & Bauer, März 2022). Die Evaluation wird als *preformative ex-ante Evaluation* eingestuft, da sie Informationen zur weiteren Vorhabengestaltung liefern soll.

Capacity Development und Qualitätssicherung

Schwerpunkt der dritten Phase liegt auf der *Kompetenzförderung (CD)* unterschiedlicher Zielgruppen. Über das Toolkit werden im sog. *Learning Lab* einführende Webseminare und mehrtägige Fortbildungsworkshops zu acht unterschiedlichen SfE-Themen (*Health, Education, Gender Equality, Employability, Social Cohesion, Inclusion, Environment, Violence Prevention*) sowie drei Module zum Thema M&E und Wirkungsmessung auf verschiedenen Sprachen (Englisch, Französisch, Spanisch, Arabisch) angeboten. Daten zu den CD-Maßnahmen werden kontinuierlich gesammelt (Monitoring) und einmal jährlich i.R. der Berichtserstattung für die GIZ ausgewertet (*formativ-summative on-going Evaluation*).

Durch die weitere Standardisierung von SfE-Materialien, -Lehr- und -Lernprozessen findet eine *Qualitätssicherung und Weiterentwicklung* der SfE-Produkte der GIZ statt. Die Entwicklung von Leitfäden wie etwa zum Reflexionsprozess, der für den Transfer von im Training erlebter und erlernter Kompetenzen in den Alltag unabdingbar ist, von exemplarischen SfE-Aktivitäten und Trainingseinheiten zu unterschiedlichen SDG-Schwerpunkten und die Weiterentwicklung von bspw. Kompetenzrastern und Inhalten des S4D Resource Toolkit sind Gegenstände dieses Arbeitsbereiches. Alle Standards werden über das Toolkit veröffentlicht.

6.5 Zusammenfassung und Bewertung der Evaluationsprozesse

Tabelle 7 fasst alle Evaluationen der drei Phasen zusammen:

Tab. 7: Durchgeführte Evaluationen Phase I-III (eigene Darstellung)

Phase	Beschreibung der Evaluation(en)	Perspektive	Befragte, Land/Länder	Methode(n)
1	1 Analyse internationaler Akteur*innen	ex-ante, preformativ	Akteur*innen SfE	Dokumentenanalyse und Einzelinterviews
1	1 Evaluation bestehender, abgeschlossener SfE-Maßnahmen	ex-post, preformativ	Mitarbeitende, Partnerorganisationen ausgewählter SfE-Projekte	Dokumentenanalyse und Einzelinterviews
3	1 Evaluation („Mapping“) zu Hochschulen	ex-ante, preformativ	Universitätsvertretende ausgewählter Länder	Dokumentenanalyse, Onlinefragebogen und Interviews
1	1 Nationale Akteursstudie	ex-ante, preformativ	18 Akteur*innen in KO	Einzelinterviews
1	1 Nationale Akteursstudie	ex-ante, preformativ	19 Akteur*innen in MO	Einzelinterviews
1, 2	38 Evaluationen i.R. der Fortbildungsmaßnahmen in den Partnerländern	on-going, formativ	Coaches, Instructors aus 6 Ländern	Prä-, Post-Fragebogen oder FGD
1, 2	2 Evaluationen zu Follow-up Besuchen bei ausgebildeten Coaches ³⁴¹	on-going, formativ-summativ	132 Coaches aus fünf Ländern	Beobachtung und Einzelinterviews
2	1 Online-Befragung	on-going, formativ-summativ	686 Coaches aus fünf Ländern; 102 Instructors aus zwei Ländern	Fragebogen
2	1 Befragung der Partnerorganisationen (Scaling-up)	on-going, formativ-summativ	14 Partnerorganisationen aus sieben Ländern	Fragebogen
1	1 Wirkungsevaluation Projekte Promundo	on-going, summativ	Kinder, Jugendliche aus BR	Prä-, Post-Fragebogen und FGD
1	1 Wirkungsevaluation i.R. von Jugendcamps	on-going, summativ	Jugendliche aus PG	Prä-, Post-Fragebogen
1	1 Wirkungsevaluation	on-going, summativ	Kinder, Jugendliche aus NA	Prä-, Post-Fragebogen
1	1 Wirkungsevaluation	on-going, summativ	Kinder, Jugendliche aus NA	Interviews oder FGD
2	1 Wirkungsevaluation bei vier ausgewählten NGOs	on-going, summativ	Kinder, Jugendliche aus BR	Prä-, Post-Fragebogen
3	5 Wirkungsevaluationen	on-going, summativ	Kinder und Jugendliche	Prä-, Post-Fragebogen
3	1 Evaluation zu allen durchgeführten CD-Maßnahmen	on-going, formativ-summativ	Teilnehmende CD-Maßnahmen	Feedbackbefragung und Online-Fragebogen

³⁴¹ Aufgrund der Stichprobengrößen wurden lediglich die Daten aus Afghanistan und Kolumbien komplett ausgewertet, weswegen an dieser Stelle nur zwei Evaluationen gezählt werden.

Insgesamt wurden knapp 60 Evaluationen durchgeführt, die meisten davon in der ersten Phase der wissenschaftlichen Begleitung. Die Mehrheit der Evaluationen fand während der Implementierungsphase der Projekte statt (*on-going*), vier in der Planungsphase der Maßnahmen (*ex-ante*), eine im Nachgang von Projekten (*ex-post*). Die Evaluationen in ihrem Umfang sind sehr unterschiedlich, bspw. wurden i.R. der Vorher-Nachher-Befragungen zu den Fortbildungen eher weniger Daten erhoben und Ressourcen investiert, wohingegen z.B. die Akteursstudien sehr umfangreich waren – sowohl aufgrund der Erhebung und Auswertung qualitativer Daten als auch bzgl. der Ressourcen (Zeit, Personal). Die meisten Untersuchungen wurden auf Ebene der Coaches und Instructors initiiert (mikro), weniger auf der Ebene von Organisationen oder anderen Akteur*innen (mikro und meso) (7), sechs davon auf der Ebene von Kindern und Jugendlichen. Im Rahmen der meisten Studien kamen/kommen eher quantitativ angelegte Fragebögen zum Einsatz, bei einigen rein qualitative Methoden (Interviews) oder Mixed-Methods-Ansätze. Zur *Beteiligung von Akteur*innen im Evaluationsprozess* kann festgehalten werden, dass die DSHS (*,extern^*) bei fast allen Evaluationen in die Konzipierung, die Entwicklung von Designs und Instrumenten, Erhebungs- und Auswertungsprozessen von Daten sowie die Berichterstattung eingebunden war/ist. GIZ-Mitarbeitende (*,intern^*) waren/sind ebenfalls an fast allen Studien beteiligt: Zumeist in organisatorischen Rollen, etwa wenn es um die Organisation der Studien vor Ort geht, aber auch bei der Entscheidung über befragte Zielgruppen, Designs und der Entwicklung von Instrumenten waren/sind GIZ-Mitarbeitende involviert. Partnerorganisationen waren/sind hauptsächlich i.R. der Wirkungsevaluationen eingebunden, häufig bei der Entwicklung von Instrumenten und der Sammlung von Daten.

Evaluationsstandards wie jene der DeGEval (2016), an denen sich die GIZ unter anderem orientiert, wurden i.R. der wissenschaftlichen Begleitung nicht explizit als Referenz eingeführt. Im Folgenden wird reflektiert, inwiefern auf ausgewählte Standardbereiche eingegangen wurde.³⁴² Aus vertraglichen Gründen kann dabei nicht auf alle Standards eingegangen werden.

³⁴² Eine ausführliche Reflexion zu allen Standards im Sinne einer *Meta-Evaluation* kann an dieser Stelle aufgrund allgemeiner Vertragsbedingungen (vgl. GIZ, Juli 2018) und aus Gründen der persönlichen Involvement nicht geleistet werden.

Gemäß den *Nützlichkeitsstandards* wurde i.d.R. sichergestellt, dass Beteiligte vorab identifiziert wurden (N1) und die Evaluationen bestimmte und geklärte Zwecke verfolgten (N2). Bzgl. des Informationsumfangs wurden v.a. in der ersten Phase i.R. von Evaluationen noch zu viele Informationen gesammelt, die nicht genutzt wurden. Andererseits fehlten wichtige Informationen lange, da entscheidende M&E-Prozesse noch nicht standardisiert waren, was sich im Laufe der Zeit jedoch regulierte (N4). Bzgl. der *Berichterstattung*³⁴³ standen die wissenschaftlichen Kooperationspartner*innen vor der Herausforderung, komplexe Ergebnisse in komprimierter Form darzustellen und zeitgleich dem wissenschaftlichen Anspruch der Vollständigkeit (u.a. Abbildung aller Ergebnisse, d.h. positive und negative) gerecht zu werden. Die Berichtsformate waren häufig zu komplex und umfangreich (N6).

Bzgl. der *Durchführbarkeit* von Evaluationen mussten Evaluationsverfahren in den Partnerländern den zur Verfügung stehenden Ressourcen und an die jeweiligen Projektkontexte angepasst werden, wodurch i.R. einzelner Teilprozesse Kompromisse (bspw. bei der Wahl von Methoden, Zeitpunkt der Datenerhebung, Umfang von Studien etc.) eingegangen wurden (D1). Um ein größeres Verständnis und eine Akzeptanz im Hinblick auf Evaluationen zu schaffen, wurden in einigen Ländern CD-Maßnahmen mit den Partnerorganisationen durchgeführt (D2).

Im Sinne der *Fairnessstandards* wurde versucht sicherzustellen, dass i.R. der Evaluationsprozesse respektvoll und fair mit beteiligten Akteur*innen umgegangen wurde. In einzelnen Fällen wurde jedoch versäumt, Rechte und Pflichten beteiligter NGOs schriftlich festzuhalten, wodurch bspw. Daten zu spät oder nicht übermittelt wurden (F1). I.d.R. wurden Evaluationen so geplant und durchgeführt, dass Rechte, Sicherheit und Würde von einbezogenen Personen geschützt waren, bspw. durch die Anonymisierung von Daten (F2). Über die Evaluationen wurde stets versucht, die Stärken und Schwächen der Evaluationsgegenstände fair und umfassend zu prüfen und bei der Berichterstattung

³⁴³ Damit Erwartungen „beider Seiten“, d.h. der Praxis und Wissenschaft, nicht enttäuscht werden, müssen Evaluationsergebnisse, die seitens der Wissenschaft dem Auftraggebenden zur Verfügung gestellt werden (1) *entscheidungsrelevante* Informationen liefern, (2) (bzgl. der Entscheidungsfindung) *zeitgerecht* geliefert werden und (3) nicht zu *komplex* sein (Stockmann & Meyer, 2017b).

darzustellen. Problematisch war hierbei zum einen repräsentative Stichproben zu gewinnen, die Sachverhalte umfassend dargestellt hätten und zum anderen die Erfassung von Schwächen, die seitens der Befragten möglicherweise ‚beschönigt‘ wurden (*Störeffekt*) (F3). An den meisten Evaluationen waren lediglich die Auftraggebenden und das Evaluationsteam beteiligt. Die Sichtweisen weiterer Beteiligten und Betroffenen (bspw. Partnerorganisationen, Coaches) auf Evaluationsgegenstände (bspw. Fortbildungs- oder Scaling-up-Maßnahmen) war durch die Befragung derselben einbezogen. Seitens der wissenschaftlichen Kooperationspartner*innen wurde versucht, den Evaluationsprozess und -berichte so unparteiisch wie möglich zu gestalten (F5). Alle Evaluationsergebnisse und -berichte standen allen Beteiligten auf GIZ-Seite über das interne Datenmanagementsystem zur Verfügung (F5).

Im Sinne der *Genauigkeit* konnten die meisten Standards eingehalten werden: Konzepte der Evaluationsgegenstände und deren Umsetzung (G1), Zwecke und Vorgehen (einschließlich methodischer Vorgehensweisen) (G3) sowie genutzte Informationsquellen (G4) wurden i.R. der Evaluationsberichte beschrieben und dokumentiert. Wenn möglich wurden Kontexte der Evaluationsgegenstände erörtert und bei der Ergebnisinterpretation berücksichtigt (G2). Gesammelte, aufbereitete, analysierte und präsentierte Informationen wurden insoweit auf Fehler überprüft, wie dies möglich war (G6)³⁴⁴. Quantitative und qualitative Informationen wurden mittels inhaltsanalytischer und statistischer Verfahren angemessen und systematisch analysiert (G7). Bewertungen und Schlussfolgerungen basierten auf den ausgewerteten Ergebnissen und auf gemachten Erfahrungen im Prozess (G8). Erhebungsverfahren und Datenquellen wurden möglichst so gewählt, dass sie valide und reliable Informationen zuließen (G5). Dabei konnte die Wahl der Informationsquellen nicht immer beeinflusst werden, sondern wurde von Auftraggebenden gesteuert. Wenn mehr Ressourcen zur Verfügung gestanden hätten, hätten Evaluationen ausgeweitet (bspw. im Falle der Follow-up Besuche bei ausgebildeten Personen) oder mehrperspektivischer gearbeitet werden können, d.h. die Nutzung zusätzlicher Informationsquellen und häufiger ergänzender Erhebungsverfahren im Sinne

³⁴⁴ Das Evaluationsteam musste sich bzgl. mancher Evaluationsschritte, bspw. bei der Übersetzung, auf Partner*innen oder Unterauftragnehmende verlassen.

eines gemischtmethodischen Ansatzes. Bestanden Zweifel bzgl. der Aussagekraft der Ergebnisse, wurden diese mit den Auftraggebenden besprochen und im Bericht dokumentiert.³⁴⁵ Alle Evaluationen wurden insoweit dokumentiert, dass sie einer Meta-Evaluation unterzogen werden könnten (G9).

³⁴⁵ Bspw. bzgl. der fraglichen Repräsentativität bei Stichproben.

7. Diskussion: Zwischen Verwissenschaftlichung und Anwendungsorientierung

Im folgenden Kapitel der Arbeit werden das *Forschungsfeld SfE* und *Forschende im Kontext von SfE* vor dem Hintergrund des Spannungsfeldes zwischen Verwissenschaftlichung und Anwendungsorientierung diskutiert. Dabei gliedern und leiten die in der theoretisch-konzeptionellen Rahmung festgelegten Kriterien den Diskussionsverlauf. Durch die Gegenüberstellung der Evaluationsforschung in der ‚allgemeinen EZ‘ und der ‚sport-spezifischen EZ‘ sowie unter Einbezug der gemachten Erfahrungen i.R. der wissenschaftlichen Begleitung des SV SfE der GIZ werden Gemeinsamkeiten, Unterschiede und Besonderheiten der Bereiche bzgl. der Untersuchungsperspektiven herausgearbeitet. Letztendlich werden Schlussfolgerungen und Ausblicke formuliert, die zum einen das Forschungsfeld besser verorten und zum anderen SfE-Forschende dabei unterstützen können, sich im Spannungsfeld zu positionieren. Das Kapitel schließt mit einer Reflexion zur theoretisch-konzeptionellen Rahmung, zur methodischen Vorgehensweise und zur eigenen Rolle im Hinblick auf diese Arbeit.

7.1 Das Forschungsfeld SfE und SfE-Forschende

7.1.1 Entwicklung: Parallelen zur EZ-Evaluation – zeitlich verzögert und wenig genutzt

Die chronologische Abhandlung zur *Evaluation in der allgemeinen EZ* (vgl. Kap. 4) zeigt deutlich, dass die Entwicklung der Evaluationsforschung eng mit entwicklungspolitischen Ereignissen und den damit verbundenen Schwerpunkten in der Entwicklungszusammenarbeit verbunden ist (Kjekhus, 1991; Neubert, 2016; Stockmann, Menzel & Nuscheler, 2016). Entwicklungspolitische Kontexte und Meilensteine hatten und haben einen Einfluss auf die Entstehungsgeschichte, auf thematische Schwerpunkte, Forschungsfragen und -gegenstände i.R. von Studien, Ansätzen und Methoden sowie auf Standardisierungs-, Institutionalisierungs- und Professionalisierungsprozesse. Die Ausführungen zur *Entwicklung des Forschungsfeldes SfE* (vgl. Kap. 5.3) veranschaulichen, dass die (Evaluations-)Forschung in der sportbezogenen EZ ähnliche Entwicklungen durchläuft, bspw. i.R. von Methodendiskussionen und Debatten um Evaluationstheo-

rien, allerdings *zeitverzögert* ab den 2000ern, da ab diesem Zeitpunkt der Sport zunehmend als Medium zur Erreichung entwicklungspolitischer Zielsetzungen an Akzeptanz erfuhr. Einflussreiche Meilensteine und entwicklungspolitische Kontexte sind selbstverständlich dieselben wie in der allgemeinen EZ (MDGs, Paris Deklaration, Agenda 2030 und SDGs), der Sport wurde jedoch im Entwicklungskontext von den Vereinten Nationen bald als „Allzweckwaffe“ und als „low-cost high-impact tool“ bezeichnet (Brand, 2020, S. 113; Kay & Dudfield, 2013, S. 5). Dies steigerte den *Druck* auf die Forschung, diesen ‚außerordentlichen Beitrag‘ von Sport nachzuweisen immens, was sich in dem sprunghaften Anstieg von SfE-Studien ab Mitte der Nullerjahre zeigt. Ähnlich wie in der EZ-Evaluation hatten anfänglich die Vereinten Nationen, später das Commonwealth Sekretariat eine führende Rolle bei Institutionalisierungs-, Professionalisierungs- und Standardisierungsprozessen, und auch die Bearbeitung durch Beratungsfirmen nahm auf ähnliche Weise zu. Was im Vergleich zur allgemeinen EZ-Evaluation im Hinblick auf die Institutionalisierungs- und Professionalisierungsprozesse eindeutig fehlt, ist jegliche Form von internationalen *Zusammenschlüssen oder Netzwerkstrukturen*, die sich auf die sportspezifische Evaluationsforschung spezialisiert hätten. Austauschformate wie eine *Next Step Conference*, bei denen auch Forschende beteiligt waren, gibt es nicht mehr. Lediglich die Online-Plattform *sportanddev.org*, das *Journal of Sport for Development (JSFD)* sowie einzelne Akteur*innen (Laureus, Streetfootballworld) nehmen sich dem Thema an und geben Einblicke in aktuelle Studien und Forschungsthemen. Das Forschungsfeld wird vornehmlich von Forschenden oder Evaluierenden aus dem Globalen Norden geprägt: Jene aus den USA, UK, Australien und Kanada publizieren die meisten Artikel. Außer aus Südafrika stammen kaum weitere Forschende aus dem Globalen Süden. Deutschland steht in diesem Vergleich weit hinten (vgl. Schulenkorf et al., 2016).

Dass Sport in der deutschen Entwicklungspolitik bereits ab Mitte der 1970er präsent war, ist international nicht bekannt (s. Kap. 2). Auch hier setzte das Forschungsinteresse verzögert ab Mitte der 1980er seitens vereinzelter Personen ein, die in sportbezogene EZ-Maßnahmen involviert waren (z.B. Giebenhain, Trosien) oder mit Studien zur sportbezogenen EZ durch das BMZ beauftragt wurden (Digel und Fornoff). Der Grund, dass die SfE-Maßnahmen deutscher Akteur*innen international keine Erwähnung finden, liegt vermutlich an den ausschließlich deutschsprachigen Veröffentlichungen. Da der

Sport ab ca. den 1990ern wieder aus den BMZ-Maßnahmen verschwand (vgl. Kap. 2.1), befassten sich auch immer weniger deutsche Publikationen mit der Thematik – ein Indiz, dass Forschungsaktivitäten eng an das entwicklungspolitische Interesse an Sport gekoppelt und von dadurch bereitgestellten Geldern abhängig sind. Durch internationale Bewegungen motiviert, erscheint 2011 ein deutschsprachiger Sammelband mit theoretischen und praktischen Beiträgen (Petry et al., 2011). Die hauptsächlichen EZ- Akteur*innen werden kurz erläutert (vgl. Groll & Hillbrig, 2011), praktische Beiträge erfolgen nur durch das AA (Steinmann, 2011; Wnendt, 2011), keiner durch das BMZ oder die GTZ. Trotz des sprunghaften Anstiegs der BMZ-Fördermittel (vgl. Abb. 1), der Etablierung des SV (s. Kap. 2.2) und der umfangreichen wissenschaftlichen Begleitung durch die DSHS ab 2013 (s. Kap. 6) erscheint erst 2020 ein weiterer Sammelband (Petry et al., 2020), diesmal auch mit einem Vorwort des BMZ und Beiträgen der GIZ. Aus Sicht der Autorin ist es ein *Versäumnis*, dass außer einer vereinzelt Publikation der Universitäten Bielefeld und Göttingen (Schreiner et al., 2020) keine weiteren Beiträge international bzw. in englischer Sprache veröffentlicht wurden.

Im *Hochschulkontext* kämpft das Forschungsfeld immer noch um Anerkennung, auch wenn es vermehrt i.R. von Teilmodulen in Masterstudiengängen verankert ist. Zusammenfassend kann festgehalten werden, dass die Bemühungen um mehr wissenschaftliche Anerkennung ‚zweigleisig‘ fahren: Zum einen versucht sich SfE i.R. neuer akademischer Programme und eigener Journals als eigenständiges Forschungsfeld zu etablieren, zum anderen wird sich um die Anschlussfähigkeit und Verortung in bereits etablierten Disziplinen wie bspw. der Entwicklungsforschung bemüht.

Es ist selbstverständlich, dass sich Evaluationsforschung im Allgemeinen, die sich über 60 Jahre lang entwickelt hat, nicht direkt mit einem jungen, themenspezifischen Forschungsfeld vergleichen lassen kann. Doch gerade wegen dieser langjährigen Entwicklungsgeschichte im EZ-Kontext und der Nähe des SfE-Kontexts zur EZ, sollten die zahlreichen evaluationsspezifischen Kenntnisse *intensiver genutzt* werden, was bislang nur wenig passiert.

7.1.2 Definitionen: Diskurstiftende Ungenauigkeiten

Dass langjährige und umfangreich vorhandene Kenntnisse aus der Evaluationsforschung in der allgemeinen EZ im SfE-Kontext wenig Beachtung finden, zeigt sich auch in der Nutzung von Begrifflichkeiten. Die Ausführungen unter Kapitel 5.1 veranschaulichen, dass im SfE-Kontext Uneinigkeit herrscht und begrifflich häufig ‚unsauber‘ gearbeitet wird, d.h. die Begriffe *research*, *evaluation* und *monitoring* werden nicht deutlich genug voneinander abgegrenzt und alte (*M&E*) sowie neue Wortverschmelzungen (*MEL*) werden abwechslungsreich genutzt. Es kann festgehalten werden, dass diese *Unklarheit bzgl. der Definitionen* unweigerlich Auswirkung auf ganz unterschiedliche Ebenen hat und *Ursprung für zahlreiche Diskussionen* in der SfE-Forschung ist, bspw. zur Rolle von Evaluierenden und Wissenschaftler*innen (Wer forscht/evaluiert unter welchen (Vertrags-)Bedingungen, d.h. beauftragt oder aus eigenem wissenschaftlichen Antrieb?), bzgl. der Vorgehensweise (Wie wird geforscht/evaluiert, d.h. welche Designs und Methoden werden genutzt?) und der Zielsetzungen und Zweckbestimmung (Warum wird geforscht/evaluiert, d.h. zur Wissensgenerierung oder aufgrund anderer Funktionen?). Zudem spiegelt sich die *Differenzierung* zwischen *professionellen Evaluationen* und *Alltags- oder Laienevaluationen*³⁴⁶, aber auch die *Forschungsbetonung* wie sie in der allgemeinen deutschsprachigen Literatur häufig zu finden ist (Döring & Bortz, 2016; Stockmann, 2007a) in der englischsprachigen SfE-Literatur nicht wider: Differenzierende oder kombinierte Begriffe, die dies betonen würden, wie bspw. *scientific evaluation* oder *evaluation research* (vgl. Vedung, 2000), sind nur selten zu finden. Es kann geschlussfolgert werden: Wenn die SfE-Forschergemeinde diese begrifflichen Unklarheiten nicht überwindet, können auch keine Fortschritte in den unterschiedlichen Debatten erfolgen. Denn je nachdem, ob Grundlagen- oder Evaluationsforschung betrieben wird, unterscheiden sich die Rollen der Forschenden, ihre Zielsetzungen und möglicherweise auch ihre Vorgehensweisen. Unweigerlich hat beides seine Daseinsberechtigung oder anders

³⁴⁶ Evaluation ist nicht gleich Evaluation: Bei *Alltags- oder Laienevaluationen* wird „[...] irgendetwas von irgendjemanden nach irgendwelchen Kriterien in irgendeiner Weise bewertet“, wodurch Evaluationen zu einer *Gefahrenquelle* werden können, indem Sachverhalte bspw. einseitig oder falsch dargestellt werden (Stockmann, 2007b, S. 27; vgl. auch Döring & Bortz, 2016, S. 992ff.). Hierbei spielt die *Qualifizierung von Evaluierenden* eine große Rolle; ein wichtiger Aspekt, der in der Definition nach Döring und Bortz (2016, S. 979) Erwähnung findet.

gesagt, beides muss nebeneinander Bestand haben. Ohne diese ausschlaggebenden Unterscheidungen wird das Forschungsfeld SfE allerdings möglicherweise in undifferenzierten Debatten ‚stecken bleiben‘ und sich nicht weiterentwickeln.

Im Rahmen der wissenschaftlichen Begleitung des SV SfE wurde vor allem der Begriff von *Wirkungen* oder *Wirkungsevaluationen* umfangreich diskutiert. Auch wenn bei Wirkungsevaluationen die mittelfristigen, direkten sowie die langfristigen Wirkungen (*Impacts*) einer Maßnahme im Mittelpunkt stehen, die nur nach einer bestimmten Zeit und unter bestimmten Voraussetzungen gemessen werden können (vgl. Kapitel 6.3.1), wurde bereits nach einer sehr kurzen Laufzeit (2015) der Versuch unternommen, Wirkungsevaluationen in den Ländern auf Ebene der Zielgruppe der Kinder und Jugendlichen durchzuführen. Neben den unterschiedlichen Auffassungen, ab wann von Wirkungen gesprochen werden kann, besteht die begriffliche Problematik auch darin, wer als *Zielgruppe* des SV SfE definiert und damit als Untersuchungsgruppe befragt wird. Eigentlich sind die Partnerorganisationen und die SfE Coaches und Instructors, mit denen Beratungs- und Trainingsmaßnahmen durchgeführt werden, *direkte Zielgruppe* des SV SfE. Kinder und Jugendliche, die über ausgebildete Coaches erreicht werden, sind eine *indirekte Zielgruppe* der Maßnahmen und sollten demnach nicht vorrangiger ‚Gegenstand‘ der Wirkungsevaluationen sein.

7.1.3 Theoretische Fundierung, Ansätze und Methoden: alle nutzen, was sie kennen – wenige nutzen, was vorhanden ist

Wie die fünfphasige Darstellung verdeutlicht (vgl. Kap. 4.1-4.5), haben sich im Verlauf der Zeit in der EZ Ansätze und Methoden entwickelt und herauskristallisiert, die mittlerweile weltweit von EZ-Maßnahmen durchführenden Institutionen genutzt und modifiziert werden, so zum Beispiel der *theoriegeleitete RBM-Ansatz* oder *partizipative Ansätze* wie der *Participatory Rural Appraisal (PRA)* (vgl. Binnendijk, 2000; Bliss et al., 2007; Oberndörfer et al., 2010). Durch die Veröffentlichungen der Stabsstelle Evaluierung der GIZ (GIZ Stabsstelle Evaluierung, 2018a-e), dem Managementhandbuch *Capacity Works* (GIZ, 2015a) und weiteren Leitfäden (BMZ, 2006, 2021) ist für die dt. EZ geregelt, welche Evaluationen und auf welche Art und Weise M&E-Aktivitäten als Teil des Projektmanagementprozesses durchgeführt werden sollen – wie bspw. das

Wirkungsorientierte Monitoring (WoM) oder die Nutzung *systemischer Wirkungsmodelle* als theoretische Grundlage. Zwar wird im allgemeinen EZ-Kontext nach wie vor die Nutzung rein quantifizierter Ergebnisse diskutiert und kritisiert (Holzapfel, 2014), jedoch ist die Nutzung von *Mixed-Methods-Ansätzen*, also der Verwendung von quantitativen und qualitativen Methoden der empirischen Sozialforschung, in der internationalen Evaluationsgemeinschaft weitgehend etabliert, so auch in der internationalen EZ-Gemeinschaft und nationalen EZ-Agenturen wie der GIZ (vgl. GIZ Stabsstelle Evaluierung, 2018a, S. 10). Auch i.R. der wissenschaftlichen Begleitung wurde auf bereits bestehende Tools der dt. EZ (bspw. Akteurslandkarten, SWOT-Analysen aus Capacity Works) zurückgegriffen oder waren Gestaltungsgegenstand (WoM). Eine Herausforderung bestand in der reinen *Quantifizierung* des onlinebasierten WoM-Systems. Qualitativ formulierte Indikatoren ohne numerische Zielwerte waren nicht oder sehr umständlich zu fassen. Darüber hinaus plädierten die wissenschaftlichen Kooperationspartner*innen für möglichst gemischt-methodische Ansätze, die allerdings aufgrund mangelnder Ressourcen nicht immer umgesetzt werden konnten.

Anders als es durch Vorgaben im allgemeinen EZ-Kontext möglich ist, wird im sportbezogenen Kontext weit weniger einheitlich evaluiert. Dies liegt vor allem darin begründet, dass SfE nicht nur als EZ-Gegenstand wahrgenommen und von EZ-Akteur*innen evaluiert wird, sondern zahlreiche Wissenschaftler*innen aus den unterschiedlichsten Disziplinen am ‚Forschungsfeld SfE‘ beteiligt sind, u.a. aus der Soziologie, Pädagogik, Psychologie, Anthropologie, den Politikwissenschaften und aus dem Bereich des Sportmanagements. Der Bereich SfE beinhaltet in seiner Komplexität eine Variation an Phänomenen, die zu Untersuchungsgegenständen gemacht werden – von Individuen auf einer Mikroebene, die im Zusammenspiel und in ihren Verhaltensweisen betrachtet werden können, über Akteur*innen und deren Organisationsformen auf einer Mesoebene bis hin zu gesellschaftlichen Auswirkungen auf der Makroebene. Zur Erklärung der Phänomene werden nach dem Motto *„jede*r nutzt, was er*sie kennt“* die unterschiedlichsten Theorien zur Fundierung von Studien verwendet, wie die Ausführungen in Kapitel 5.4.3 zeigen. Zwar sprechen sich einzelne Forschende dafür aus, verschiedene theoretische Mo-

delle zusammenzuführen (vgl. Burnett, 2015, S. 388), jedoch findet aufgrund der *Multi-disziplinarität* immer noch *wenig Austausch* statt. Bewährte Modelle aus der EZ, darunter das RBM, finden sich in Coalters (2013b) Programmtheorie, die neben anderen Ansätzen als Schlüsselkonzept bezeichnet wird (Welty Peachey et al., 2019), zwar wieder, allerdings wird es wie andere Schlüsselkonzepte (vgl. Kap. 5.4.3) auch bislang wenig in weiteren Forschungsarbeiten angewendet oder weiterentwickelt (ebd.). Bereits Anfang der letzten Dekade verwiesen Autor*innen auf die Dringlichkeit, sich eingehender mit *Literatur aus der allgemeinen Entwicklungs- und EZ-Forschung* auseinanderzusetzen und diese auch zu nutzen (vgl. Levermore, 2011; Darnell & Black, 2011; Darnell, 2012; Donnelly et al., 2011), in den letzten Jahren werden es immer mehr (vgl. Darnell & Dao, 2017, S. 26; Giulianotti et al., 2019, S. 411; Lindsay, 2017, S. 814; Lindsey & Bitugu, 2019, S. 88; Spaaij et al., 2018, S. 26). Dies wurde von Anfang an damit begründet, die als „trial and error“ bezeichneten Evaluationssysteme im SfE-Kontext zu verbessern und sich an bestehender EZ-Evaluationspraxis zu orientieren: „Despite – or rather because of – these different approaches, sport-for-development attracts criticism for its unrealistic and varied approach to evaluation“ (Levermore, 2011, S. 340). Vor allem, weil der Sport zunehmend mit EZ-Institutionen verknüpft ist (vgl. Darnell, 2012, S. 4), auf entwicklungspolitische Veränderungen reagiert werden muss (vgl. Giulianotti et al., 2019, S. 431) und eine Anschlussfähigkeit und Glaubwürdigkeit im Hinblick auf andere Disziplinen gewährleistet werden sollte (vgl. Darnell & Dao, 2015, S. 26; Welty Peachey & Cohen, 2015, S. 24), ist die Auseinandersetzung mit und die Anwendung von bewährten theoretischen Modellen aus dem EZ-Kontext grundsätzlich unabdingbar. Im Speziellen verweisen Forschende auf die Entwicklung von Diskursen³⁴⁷, von Definitionen und Evaluationsnetzwerken (vgl. Levermore & Beacom, 2012, S. 132f.), auf den Umgang mit und die Umsetzung von Wirkungsstudien (vgl. Kay & Dudfield, 2013, S. 17) sowie auf „mainstream ideas“ (Darnell & Dao, 2017, S. 26) und weitere theoretische Modelle in spezifischen Sektoren (bspw. Edwards, 2015 zu *public health*) hin. V.a. in den Veröffentlichungen von Iain Lindsey, der sich bereits früh mit der Nutzung von Literatur aus anderen EZ-Sektoren

³⁴⁷ Levermore und Beacom (2012, S. 133) verweisen auf „alternative evaluative discourses“ und nennen die Publikation des *Center for Global Development* „When will we ever learn?“ (Savedoff et al., 2006).

beschäftigt und auf SfE übertragen hat,³⁴⁸ wird die mangelhafte Auseinandersetzung mit theoretischen EZ-Konzepten deutlich: Er ist der Ansicht, dass *Probleme projektbasierter Ansätze*, die im SfE-Kontext nach wie vor noch gang und gäbe sind, ein *Spiegelbild* jener darstellen, wie sie bereits im EZ-Kontext in den 1990ern identifiziert wurden (ebd., S. 802). Diese seien geografisch, zielgruppenspezifisch, zeitlich begrenzt, gebergesteuert und unrealistisch bzgl. der Erwartungen hinsichtlich ihrer Wirkungen und Nachhaltigkeit. Ineffizienz, Fragmentierung, Duplikation von Projekten und unnötige Konkurrenz zwischen Akteur*innen sind Folgen der Ansätze (Lindsey, 2017).³⁴⁹ Weitere Beispiele für die Nutzung von EZ-Literatur sind bei Levermore (2011), Coalter (2008, 2013a) und Kay (2012) zu finden.³⁵⁰ Doch auch, wenn die Stimmen zahlreicher SfE-Autor*innen hierzu lauter werden, gehen diesem Ruf bislang noch (zu) wenige Forschende nach. Durch die Konzeptionierung von „Sportland“ läuten namenhafte SfE-Forschende³⁵¹ eine „*new research agenda*“ für den SfE-Sektor ein (vgl. Giulianotti et al., 2019).³⁵² Im Rahmen der Agenda werden „Sportland scholars“ dazu angehalten, wesentliche Änderungen im globalen Entwicklungskontext³⁵³ zu beachten und sich mit „fresh literature“ aus dem Bereich zu befassen (ebd., S. 411). Die „Theoretisierung“ sollte nach Meinung der Autor*innen *zeitgemäß, kritisch und partizipativ* sein. Es sollen zum einen Ansätze wie *capabilities* oder *societal transformation approach* genutzt werden, die die Mehrdimensionalität von Entwicklung fassen und die Lücke zwischen „small“ und „big development“ schließen, und zum anderen kernpolitische Zwickmühlen (bspw. Gebende-Empfan-

³⁴⁸ Bspw. zur Entwicklung eines *Nachhaltigkeitskonzepts* (Lindsey, 2008), zu *Partnerschaften* (vgl. Lindsey & Bitugu, 2019) und der Notwendigkeit von Konzepten, die die komplexen *Beziehungen zwischen staatlichen und nichtstaatlichen Akteur*innen* veranschaulichen (Lindsey, Chapman & Dudfield, 2020) sowie zum Thema *Governance* (Lindsey, 2017).

³⁴⁹ In diesem Zusammenhang verweist er auf *Sector Wide Approaches (SWAPs)*, die in der allgemeinen EZ ab Mitte der 1990er Einzug fanden. SWAPs werden lediglich einmalig im Dokument *Harnessing the Power of Sport for Development and Peace* (SDP IWG, 2008) in einem Satz erwähnt (Lindsey, 2017).

³⁵⁰ Levermore (2011) nutzt u.a. Literatur von Cracknell (2000) zur Erläuterung von gängigen Evaluationsmethoden und -ansätzen im SfE-Kontext. Coalter (2008, 2013b) nutzt i.R. seiner Programmtheorie umfangreich die Beiträge zur Politikevaluation von Weiss (1993; 1997). Kay (2012) fundiert ihre Kritik an der SfE-Evaluationspraxis einschlägig auf Beiträgen des Sammelwerks *Inclusive Aid, Changing Power and Relationships in International Development* (Groves & Hinton, 2004).

³⁵¹ Richard Giulianotti, Fred Coalter, Holly Collison und Simon C. Darnell.

³⁵² Der Begriff „Sportland“ ist hierbei von den Konzepten „Aidland“ und „Peaceland“ aus dem allgemeinen Entwicklungskontext abgeleitet und betont dadurch die enge Verbindung von Sport zu diesem.

³⁵³ Bspw. zunehmender Nationalismus, aufstrebende Staaten, globale Ungleichheiten, die über die SDGs thematisiert werden, neue Kooperationsmodelle oder die Abkehr vom Neoliberalismus (ebd., S. 416ff).

gende Bevormundung) und theoretisch-praktische Dilemmata (bspw. Spannungsverhältnis zwischen Wissenschaft und praktischer Evaluation) Beachtung finden (ebd., S. 419f.). Der Artikel führt bereits bestehende Debatten zusammen und vereinigt notwendige Schritte in einer *neuen Forschungsagenda*. Es bleibt abzuwarten, wie diese seitens der SfE-Forschergemeinde aufgenommen und umgesetzt wird.

Im EZ-Kontext sind Reviews, u.a. der OECD, gängige Praxis, für die Zukunft wird eine weitere Zunahme an Evaluationssynthesen vorhergesagt, bspw. um Erkenntnisse zu einem bestimmten SDG abbilden zu können. Auch im SfE-Kontext ist eine *Zunahme an systematischen Reviews und Metaanalysen* festzustellen, was als wegweisend für die strategische Einbettung von Sport in umfassendere Entwicklungsstrategien und für die methodische Weiterentwicklung des Forschungsfeldes (Nutzung Logischer Modelle, Durchführung qualitativ hochwertiger Wirkungsstudien) eingeschätzt wird (vgl. Darnell et al., 2019).

Kapitel 5.3 macht deutlich, dass Forschung im SfE-Kontext und damit verbundene Ansätze und Methoden generell unter *großem Einfluss der internationalen Gebergemeinschaft* steht, die seit den 2000ern Sport zunehmend zur Erreichung von Entwicklungszielen einsetzt (vgl. Giuliantotti et al., 2019). Veröffentlichungen und Leitfäden staatlicher EZ-Organisationen (z.B. UN, Commonwealth Secretariat) legen also Schwerpunkte fest und prägen das Wirkungsverständnis im SfE-Kontext, d.h. die Orientierung und den Beitrag zu spezifischen Entwicklungszielen. Folglich wurden SfE-Maßnahmen, seien sie initiiert durch die UN, dem Commonwealth Secretariat oder der GIZ, von Beginn an mittels der intern existenten M&E-Systeme auf ihre Zielerreichung und Erfolge ‚geprüft‘, wie jegliche weitere sportunspezifischen EZ-Maßnahmen auch. Diese am RBM orientierten Systeme sind schwerwiegender Kritik ausgesetzt: Kay (2012) ist der Ansicht, dass M&E-Systeme hierarchische Verhältnisse stützen und lediglich dazu dienen, dass Empfängerorganisationen gegenüber dem ‚Gebenden‘ Rechenschaft ablegen können:

„In a field in which power relationships are so problematic, M&E systems are thus the most visible mechanisms for making ‘recipients’ of sport in development support accountable to its ‘donors’. For many, this makes M&E the most immediate and explicit embodiment of the power imbalances that permeate this policy field.” (Kay, 2012, S. 891)

In diesem Zusammenhang ist allerdings anzumerken, dass sich die Argumentation in ihrem kritischsten Artikel hauptsächlich auf *ein EZ-Sammelwerk* stützt (s. Groves & Hinton, 2004), in der u.a. lokale Projektmitarbeitende davon berichten, dass sie dazu angehalten wurden, den ZOPP-Ansatz der GIZ zu nutzen (vgl. Win, 2004). M&E-Systeme werden hierbei *in ihrer Gesamtheit*, inklusive ihrer verschiedenen Zielsetzungen (Lernen, Entwicklung) und Möglichkeiten *nicht erfasst*.

Es ist logisch, dass ein Forschungsfeld, das durch Forschende diverser Disziplinen mit verschiedenen theoretischen Ansätzen gestaltet ist, auch im Hinblick der methodischen Vorgehensweisen keine Einigkeit aufweisen kann (vgl. Kap. 5.4.4). Auf den ersten Blick ist es ‚nur‘ ein Methodenstreit wie er nach wie vor in der empirischen Sozialforschung stattfindet, d.h. eine Aufspaltung in Anhänger*innen des quantitativen und qualitativen „Lagers“.³⁵⁴ Bei näherer Betrachtung geht der Streit allerdings über die Diskussion um die „richtigen“ methodischen Vorgehensweisen hinaus und mündet in grundlegenden *Debatten zu ‚top-down‘ und ‚bottom-up‘-Ansätzen in der EZ*. Gegenstände sind Wissenschaftstraditionen, die Forschungsperspektive der Forschenden aus dem Globalen Norden (*‚colonizer‘*), die Projekte und Organisationen im Globalen Süden (*‚post-colonial‘*) evaluieren, die Auffassung und Definition von ‚Entwicklung‘ allgemein sowie der Produktion von Wissen (*‚knowledge production‘*) im Speziellen (vgl. Black, 2019, S. 36; Whitley et al., 2020, S. 22). Kritiker hinterfragen positivistische/quantitative Methoden und sind beunruhigt, dass durch diese Vorgehensweise postkoloniale Machtstrukturen und hegemonische Systeme verstärkt und zeitgleich lokale Erkenntnisgewinnung unterdrückt wird (Darnell & Hayhurst, 2012; Kay, 2012; Lindsey & Grattan, 2012; Nicholls et al., 2011). Gegenstimmen (vgl. Coalter, 2013a, S. 49) stellen dies als Kritik von militanten „Befreiungsforschern“ dar, die versuchen, neokoloniale Einstellungen mit positivistischen/quantitativen methodischen Vorgehensweisen auf eine Stufe zu stellen und es meiden, Outcomes und Impacts zu definieren und zu messen (vgl. Whitley et al., 2020, S. 22f.). Der Schlagabtausch wird mal mehr, mal weniger umfangreich über jegliche Art von Veröffentlichungen geführt (bspw. Coalter, 2013a; Darnell et al., 2016; Kay, 2009;

³⁵⁴ Zur Diskussion bzw. dem Streit um Methoden in der empirischen Sozialforschung s. zusammengefasst Caspari (2004, S. 95f.), ausführlich bspw. Homann (1989) und aktueller Flick et al. (2014). S. zusammenfassend für den SfE-Kontext auch Massey (2019, S. 177).

2012; Kay & Mansfield, 2016; Lindsey & Grattan, 2012; Saavedra, 2019). Die unterschiedlichen Positionen werden in Ausführungen zu den Schwerpunkten, genutzten Ansätzen und Methoden in Kapitel 5.4.4 deutlich und weisen Parallelen zur allgemeinen EZ-Kritik auf, auf die bereits in Kapitel 4.2 und 4.3 an einigen Stellen eingegangen wurde.³⁵⁵ Debatten, die in der Sozialforschung eine lange Tradition haben, werden von beiden Seiten wenig einbezogen und wichtige mehrperspektivische Betrachtungsweisen, bspw. zu unterschiedlichen Evaluationsfunktionen, Beziehungen zwischen Evaluierenden und Auftraggebenden, Differenzierungen zwischen Auftragsforschung und freier Wissenschaft sowie dem Verhältnis zwischen Wissenschaft und Politik, werden unzureichend reflektiert (vgl. Coalter, 2013a, S. 47). Massey und Whitley (2019, S. 177) sind der Ansicht, dass anstatt des weiteren Schlagabtausches besser eine differenziertere Methodendebatte stattfinden sollte: „[...] rather than lay blanket critiques across different research paradigms and epistemologies, there is a need to discuss higher levels of sophistication in both instrumental/ positivist (i.e. quantitative) and descriptive/ critical (i.e. qualitative) SDP research.“

Die SfE-Forschergemeinde ist sich dieser *Spaltung* in vornehmlich zweier Hauptstränge, „[...] that respectively emphasise positivistic, evidence-based approaches to improving SDP and critical recognition of structural constraints on the possibilities of SDP“ (Lindsey & Darby, 2019, S. 809), bewusst. Sie ist der Ansicht, diese überwinden zu können, etwa durch die Nutzung theoretischer Ansätze, die sich auf *systemische Zusammenhänge* fokussieren – wie bspw. das Konzept der Politikkohärenz (vgl. ebd.) oder die Akteurs-Netzwerktheorie (vgl. Darnell et al., 2018b). Letztere Theorie geht auf einen wichtigen Aspekt ein, der eine große Rolle bei der Unterscheidung dieser zwei Hauptstränge spielt und in Folgekapiteln Diskussionsgegenstand sein wird: *auf die beteiligten Akteur*innen*. Was Forschungen im SfE-Kontext im Großen und Ganzen voneinander unterscheidet, ist maßgeblich davon abhängig, *welche Akteur*innen am Forschungsprozess beteiligt sind*, ob und wie dieser von einer Auftragslage bedingt wird und welches Forschungsinteresse

³⁵⁵ Black (2017, 2019) erläutert in seinen Beiträgen ausführlich die Unterscheidung zwischen ‚top-down‘ und ‚bottom-up‘ in der allgemeinen EZ und reflektiert die Dynamiken im Bereich von SfE.

verfolgt wird.³⁵⁶ Forschende versuchen zunehmend die Lücke zwischen den Hauptströmen zu schließen und eine „Entweder-oder-Perspektive“ zu vermeiden, indem i.R. wissenschaftlicher Artikel differenzierter diskutiert (vgl. Whitley et al., 2020, S. 22) und nach theoretischen und methodischen Lösungen gesucht wird.³⁵⁷

7.1.4 Funktion und Nutzen: Evaluation ist mehr als ‚accountability‘ und ‚learning‘

Wie in Kapitel 4 beschrieben, standen in den 1980ern mit der Entwicklung von Logical-Framework-Techniken die *Kontroll- und Steuerungsfunktion* in der EZ im Vordergrund (vgl. Cracknell, 2002; Berlage & Stokke, 1992; Sasaki, 2006). Auch die erste, im BMZ eingerichtete Evaluationsabteilung wurde später als *Zentrale Erfolgskontrolle* bezeichnet (Wollmann, 2017). Ab den 1990ern rückte zunehmend die *Lernfunktion* in den Vordergrund, der Kontrollaspekt wurde bei der Befragung deutscher EZ- Akteur*innen kaum noch erwähnt (vgl. Borrmann & Stockmann, 2009a). Durch die Abbildung v.a. positiver Ergebnisse wurden Evaluationen fortan vermehrt auch zur positiven Außendarstellung genutzt (vgl. Caspari, 2006). Mit der Einführung der Agenda 2030 und dem damit verbundenen komplexen, mehrdimensionalen Zielsystem übernimmt Evaluation bis heute auch eine *zentrale Aufklärungsfunktion*, die bislang am schwächsten entwickelt war (vgl. Stockmann, 2017). Zudem geht der Trend hin zu einer *ausgeprägten Legitimationsfunktion*, d.h. Evaluationen zu Legitimationszwecken gewinnen an Bedeutung (Stockmann & Meyer, 2017). Evaluationsforschende sind sich „[...] weitgehend einig, dass Evaluationen im politischen System immer weniger zu Kontrollzwecken genutzt, sondern vermittelt über die zunehmende Wirkungsorientierung (s. hierzu Ausführungen im folgenden Teilkapitel), immer mehr zur Legitimierung politischer Maßnahmen und Programme eingesetzt werden“ (Meyer & Stockmann, 2017, S. 255). Auch die prominente Platzierung von Wirkungsevaluationen in der dritten Phase der SV nach über acht Jahren Projektimplementierung könnte als Legitimierungsprozess der SfE-Maßnahmen verstanden werden.

³⁵⁶ Bspw. müssen staatliche Akteur*innen, aber auch nichtstaatliche Organisationen Vorgaben einhalten, um ihr Investment zu legitimieren und auch ihren Fortschritt (bspw. bzgl. der SDGs) zu kommunizieren. Dies verlangt das Vorlegen aggregierbarer Daten, die intensive Monitoringsprozesse und quantifizierte Vorgehensweisen voraussetzen (vgl. Lindsey & Chapman, 2017, S. xvi, 33; Whitley et al., 2020, S. 25).

³⁵⁷ S. bspw. Darnell et al. (2018b), Giulianotti et al. (2019) und Lindsey & Darby (2019). S. Hinweise zu *Triangulation* und *Mixed Methods* bei Giulianotti et al. (2019, S. 414), Lindsey und Chapman (2017, S. 39f.), Preti (2012, S. 316).

Die GIZ (GIZ Stabstelle Evaluierung, 2018a, S. 6) definiert aktuell drei Grundfunktionen für Evaluationen: evidenzbasierte Entscheidungen, Transparenz und Rechenschaftslegung sowie organisationales Lernen. Doch auch wenn sich die Funktionen, die der Evaluationsforschung zugeschrieben werden (vgl. Kap. 3.1.3), im Laufe der Zeit ausdifferenziert haben, *reduzieren* sie internationale Quellen (Cracknell, 2000, S. 54ff.; OECD, 2010a, S. 7) häufig noch *auf zwei*: die „*Accountability and (Lessons-) Learning Function*“.

So auch im SfE-Kontext, wo die Vernachlässigung der Lern- oder Entwicklungsfunktion jedoch in der Kritik steht (vgl. s. Kap. 5.1). Hierbei wird abermals deutlich, dass die Debatten rund um die Evaluation von SfE-Maßnahmen an den beteiligten Akteur*innen, d.h. an den ‚Gebenden‘ auf der einen und den ‚Empfangenden‘ auf der anderen Seite, ausgerichtet sind. Stehen die Geldgebenden für Begriffe bzw. Evaluationsfunktionen wie *accountability* und *control* sowie der Durchführung von *summativen Evaluationen*, werden Empfängerorganisationen mit Begriffen wie *learning*, *ownership* und *formativen Evaluationen* in Verbindung gebracht (vgl. Coalter, 2009; Kay, 2012; Levermore & Beacom, 2012). Auch nach Jeanes und Lindsey (2014, S. 200) werden M&E-Prozesse dafür instrumentalisiert, SfE als Sektor zu legitimieren „by collective evidence of positive impacts“. Coalter (2013a, S. 45) sieht die eindimensionale Funktionszuschreibung bzgl. der Beteiligten als große Hürde für die Forschung:

„[...] Evangelical policy makers and many providers have a relative indifference to research and evaluation, viewing this simply in terms of producing proof for their beliefs of legitimating their funding. Both of these make objective research in this area extremely difficult and, in some quarters, unwelcome.“

Dadurch, dass viele Stakeholder lediglich die Kontroll- und Legitimitätsfunktion mit Evaluationsprozessen in Verbindung bringen, besteht die Möglichkeit, dass aus Angst Aussagen verfälscht oder nur positive Ergebnisse kommuniziert werden (vgl. Collison et al., 2016; Jeanes & Lindsey, 2014). Vor dem weiteren Hintergrund, dass im SfE-Kontext vorwiegend im Auftrag einer Geberorganisation geforscht wird, stehen i.R. der beauftragten Evaluationen vorwiegend die *Interessen der Auftraggebenden* im Vordergrund.³⁵⁸

³⁵⁸ S. hierzu Ausführungen unter Kap. 7.1.7 und 7.1.8.

Sie bestimmen letztendlich auch, wie die Evaluationen genutzt werden und welche Funktionen sie dadurch erfüllen.

Wichtig zu bedenken ist, dass „[...] Evaluation immer *nur eine unter mehreren Entscheidungsgrundlagen* darstellt, die bei politischen Entscheidungsprozessen Berücksichtigung finden“ (Brandt, 2007, S. 167; vgl. Rossi et al., 2004, S. 375). Neben rein sachbezogenen Handlungsempfehlungen, wie sie aus Evaluationsergebnissen generiert werden können, sind auch andere Faktoren im politischen Entscheidungsfindungsprozess von Bedeutung und können eine höhere Gewichtung haben. Hinsichtlich dieser Faktoren, die durch ein politisches System bestimmt sind, haben Evaluierende keinen Einfluss (Brandt, 2007). So können über Evaluationen ermittelte Empfehlungen keine oder eine geringere Rolle bei der letztendlichen Entscheidung spielen als die Erwartungen, die über ein politisches System gestellt werden (vgl. Weiss, 1993, S. 98). Wenn sie *keine Rolle* für jegliche Prozesse (Steuerung, Entscheidungen etc.) spielen, haben sie folglich auch *keine Funktion bzw. keinen Nutzen*.

7.1.5 Evaluation von Wirkungen und Nachhaltigkeit von Maßnahmen: Herausfordernd, komplex und unterfinanziert

Wie die chronologische Darstellung von Evaluation in der EZ verdeutlicht, ist die Debatte um die Wirksamkeit von EZ-Maßnahmen eine ‚alte‘: Diskussionen in und um die EZ werden seit jeher von der Frage nach Wirkungen bzw. der Wirksamkeit bestimmt. Beeinflusst von politischen Umständen und Evaluationstrends, rückte sie mal mehr und mal weniger in den Vordergrund. Die Einführung mehrdimensionaler Begriffe wie *Nachhaltigkeit* ab den 1990ern, und mehrdimensionaler Zielsysteme wie den SDGs ab 2015 machen die Evaluation nicht nur in der EZ zunehmend komplexer. Letztendlich hat sich das Wirkungs- und Nachhaltigkeitsverständnis verändert, was sich in der Evaluationspraxis widerspiegelt: Evaluationen werden mehrheitlich am Zielsystem der SDGs ausgerichtet, Staaten stellen sich i.R. freiwilliger Review-Verfahren der Überprüfung ihrer Zielerreichung und v.a. aufgrund von Zusammenschlüssen (z.B. NoNIE, 3iE, EvalSDGs) und spezialisierten Institutionen (z.B. DEval) werden mehr und methodisch geschärfte Wirkungsevaluationen durchgeführt. Dennoch ist die dt. EZ heute noch nicht in der Lage, Nach-

haltigkeit im Sinne der Agenda 2030 abzubilden (DEval, 2018), weswegen eine *Systematisierung, Vereinheitlichung und konzeptionelle Neuausrichtung der Evaluation von Nachhaltigkeit* im deutschen EZ-System längst überfällig ist, v.a. vor dem Hintergrund, dass die Evaluation nachhaltiger Entwicklung zunehmend an Bedeutung gewinnen wird (vgl. Martinuzzi & Meyer, 2017, S. 93). Zudem stellt sich weiterhin die ‚alte Frage‘ nach der Wirksamkeit von EZ, an der die breite Bevölkerung zweifelt (vgl. Schneider & Gleser, 2018). Die bereits zahlreich existierenden theoretisch-fundierten Ansätze und Modelle zur Wirkungsevaluation, die als „gut zugänglich“ eingestuft sind (vgl. Oberndörfer et al., 2010, S. 216), werden bislang noch zu wenig genutzt und von daher auch zu wenig erprobt und weiterentwickelt. Ex-post Evaluationen, die ausschließlich die Nachhaltigkeit und die langfristigen Wirkungen (Impacts) nach Abschluss eines Programms untersuchen, kommen bei der KfW zwar standardisiert zum Einsatz, es wird jedoch angenommen, dass die GIZ nach wie vor noch *sehr wenige ex-post Evaluationen* durchführt³⁵⁹ und die Bewertung der Nachhaltigkeit von Projekten oder Programmen v.a. im Sinne ihrer *Dauerhaftigkeit* vorgenommen wird. U.a. ist dies darauf zurückzuführen, dass die Planungsphase der Programmgestaltung und -umsetzung nach dem BMZ-Verfahren sehr umfangreich (fünf Schritte) ist, wohingegen der ex-post Wirkungsphase nur wenig Aufmerksamkeit geschenkt wird (vgl. BMZ, Oktober 2017, S. 7ff.). Eine Bewertung erfolgt vornehmlich in Form einer prognostizierenden Aussage i.R. von Schlussevaluationen, die am Ende eines Projekts/Programms durchgeführt werden.

Um tatsächliche Aussagen bzgl. der Wirkungen und Nachhaltigkeit der SfE-Projekte treffen zu können, wäre es i.R. der wissenschaftlichen Begleitung notwendig gewesen, Evaluationen auch eine gewisse Zeitspanne *nach Beendigung der Maßnahmen* im Land durchzuführen. Jedoch sind unter den beschriebenen Evaluationen unter Kapitel 5 nur wenige als *ex-post* bezeichnet. Der genutzte *Evaluationsansatz zum Scaling-up* erfasst die nachhaltige Verankerung der SfE-Ansätze und Methoden bei Partnerorganisationen und deren Implementierung durch ausgebildete Personen. Er wird als vielversprechende

³⁵⁹ Im Evaluationsbericht von 2020 (GIZ Stabsstelle Evaluierung, Dezember 2020) wird zwar von 39 abgeschlossenen Zentrale Projektevaluierungen (ZPE) gesprochen, allerdings geht aus den Ausführungen nicht hervor, ob diese als ex-post, Schluss- oder Zwischenevaluierungen durchgeführt wurden. Dass die „Auswertungen immer wieder [zeigen], dass die Benotungen der Nachhaltigkeit bei Zwischenevaluierungen, die diese nur abschätzen können, meist positiver ausfallen als bei Schluss- oder Ex-Post-Evaluierungen“ (ebd., S. 97) ist der einzige Hinweis auf die Unterschiedlichkeit der Evaluationen.

Möglichkeit eingeschätzt, herauszufinden, welche Scaling-up-Strategien am nachhaltigsten bzgl. der Verankerung und am wirksamsten bzgl. der Verbreitung von SfE-Inhalten sind und könnte für weitere Studien verwendet werden.

Generell sehen Expert*innen i.R. der Wirkungsevaluationen nach wie vor *zahlreiche methodische Herausforderungen*, d.h. experimentelle Designs als ‚Goldstandard‘, die die Effekte einer Maßnahme isolieren sollen, allerdings selten umsetzbar sind, sowie *praktische* (bspw. Zugang zu Kontrollgruppen für experimentelle Designs) und *ethische*. Letztere aus dem Grund, dass bspw. pädagogische Maßnahmen zur Erklärung von Kausalmechanismen so kontrolliert und standardisiert werden müssen, dass es zur Vernachlässigung von Bedürfnissen und Belangen einzelner Kinder kommen kann. Dennoch prognostizieren sie eine qualitative Verbesserung und eine Zunahme durchgeführter Wirkungsevaluationen, v.a. wegen einer steigenden Nachfrage allgemein und zur Rechenschaftslegung und Legitimierung politischen Handelns (Dahler-Larsen, 2017; Müller, 2017; Meyer & Stockmann, 2017). Wie in den Ausführungen zu Kapitel 6 geschildert, wurden methodische und organisatorische Herausforderungen auch in der Kooperation mit dem SV SfE umfangreich diskutiert und in einem Grundsatzpapier zu Wirkungsevaluationen festgehalten.

Gemeinhin sollten *interne und externe Faktoren*, die die Wirkungsmöglichkeiten von Maßnahmen bedingen, bedacht und im Speziellen i.R. von Wirkungsevaluationen analysiert werden. V.a. *interne Organisationsstrukturen* in der EZ werden als kontraproduktiv oder als ‚hemmend‘ bzgl. der Wirkungsmöglichkeiten eingeschätzt: Z.B. hat die vorgegebene *Personalrotation* zur Korruptionsprävention (vgl. GIZ, September 2021, S. 6), wodurch Mitarbeitende häufig Rollen und Zuständigkeiten wechseln, unvermeidlich einen Einfluss auf Maßnahmen, was das Expert*innenwissen und Wissensmanagement anbelangt. Oder die *Gestaltung von Maßnahmen* an sich, wobei Projekte und Programme bzgl. ihrer Laufzeit oft zu kurz angelegt sind, um Wirkungen entfalten zu können. Oder Wirkungsevaluationen werden nicht von Beginn an eingeplant, d.h. es werden in der Projektplanung nicht die nötigen Voraussetzungen geschaffen (bspw. Baselines, Bereitstellung von Ressourcen), Wirkungen nach einer gewissen Zeit messen zu können. Generell wird auch zu wenig Fokus auf die Auslaufphase oder die Nachsorge der Projekte

gelegt, die für die Nachhaltigkeit äußerst wichtig sind. Neben diesen zahlreichen internen Faktoren beeinflussen im EZ-Kontext selbstverständlich auch *externe (Umwelt-)Faktoren* die Wirkungsmöglichkeit von Maßnahmen, v.a. politische Systeme im In- und Ausland (vgl. BMZ, Oktober 2017, S. 7ff.; Lindsey, 2017; Stockmann, 2006).

Sowohl die Abhandlungen zur zweiten Phase der wissenschaftlichen Begleitung (s. Kap. 6.3) als auch jene zu SfE-spezifischen Untersuchungsschwerpunkten und -arten zeigen auf (s. Kap. 5.4), dass im SfE-Kontext zunehmend versucht wird, den *Beitrag von Sport zu den SDGs* zu präzisieren. Letztendlich soll er über die sportspezifischen Indikatoren-Frameworks des Commonwealth messbar sein. Hier sind mögliche Beiträge von SfE-Maßnahmen und -Organisationen Indikatoren der Kategorie 3 zugeordnet, die aufgrund deren Vielfältigkeit an Schwerpunkten, Kontexten und Methoden nur exemplarisch formuliert und nicht immer aggregiert werden können. Achtgegeben werden muss darauf, dass in diesem Zusammenhang nicht reines ‚Politik-Lobbying‘ betrieben wird (vgl. Coalter, 2013a, S. 45) und Veränderungen auf einer Mikro-Ebene nicht einfach kumulative Wirkungen auf einer gesellschaftlichen Ebene bedeuten. Dies bedarf mehr Studien auf einer Meso- und Makroebene, wie es seitens der Forschenden bereits mehrfach erwähnt wurde (s. Kap. 5.4.2). Das Fehlen an Studien auf höheren Ebenen ist für Darnell et al. (2018a, S. 139f.) ein Indiz dafür, dass SfE-Forschende sich bislang *zu wenig mit globalen Trends und Strategien in der internationalen EZ* auseinandersetzen. Hierfür existieren „robust research streams“, weswegen mehr SfE-Studien auf einer internationalen oder Makro-Ebene ansetzen müssten, die bspw. Themen wie *Nachhaltigkeit* beinhalten. Auch Zanotti und Stephenson (2019, S. 166) sind der Ansicht, dass zahlreiche theoretische Ansätze auf allen Ebenen („individual, organizational and national/ transnational scales“) zur Verfügung stehen, diese aber zu wenig genutzt werden.³⁶⁰ SfE-Forschende sehen die Lösung des Problems des Mikro-Makro-Paradoxons in *systemtheoretischen Ansätzen*, was dem Trend in der allgemeinen EZ-Evaluation, von linearen hin zu systemischen Betrachtungen (von Wirkungsketten hin zu systemischen Wirkungslogiken) entspricht:

³⁶⁰ Zanotti und Stephenson (2019, S. 166ff.) nennen in diesem Zusammenhang u.a. „theories of national governance“ (z.B. im Bereich Friedensentwicklung), „theories of organizational and network learning“, „network analyses“, „studies of group/organizational change“ und „social capital“.

„A systems theory approach to SDP research can lead to the study of multiple systems levels (e.g. microsystem, mesosystem, exosystem, macrosystem), across various levels of influence (e.g. individual, school, community, policy) and influencers (e.g. parents, peers, youth workers, funders, governments, corporations), and the interaction of these factors over time and within an historical context.“ (Massey & Whitley, 2019, S. 181)

Demnach kann mittels Evaluationen, die auf systemtheoretischen Ansätzen basieren, zum einen die Lücke zwischen „micro-level program successes and macro-level changes“ geschlossen werden (Massey et al., 2015 S. 18) und zum anderen diverse Forschungsströme ‚vereint‘ werden: „[...] studying SDP through systems science may allow for SDP to become a transdisciplinary field of study that unites, rather than silos, those from divergent research traditions“ (Massey & Whitley, 2019, S. 177).³⁶¹

Die Abhandlungen demonstrieren auch, dass das *Thema Nachhaltigkeit* an sich im Bereich der SfE-Evaluation und Forschung bislang *wenig explizit bearbeitet* wurde. Zwar wird die Nachhaltigkeit (im Sinne von Dauerhaftigkeit) von geberabhängigen SfE-Projekten hinterfragt und kritisiert (vgl. bspw. Reis et al., 2017; Lindsey, 2017), die ökologische Dimension im Zusammenhang mit Sport und Umweltthemen diskutiert (vgl. Darnell, 2019) sowie i.R. des theoretischen Ansatzes S4D Framework (Schulenkorf, 2012) ‚mitgedacht‘,³⁶² doch Programmevaluationen hinsichtlich ihrer Wirkungen und Nachhaltigkeit sind rar. Dazugehörige Evaluationsansätze/-rahmen sind (noch) nicht entwickelt: „[...] there is space for the development of appropriate metrics by which to measure program impacts and determine areas of efficacy, particularly in terms of sustainability“ (Sherry & Osborne 2019, S. 382). Lediglich Lindsey (2008) hat sich eingehender mit der Konzeptionierung von Nachhaltigkeit im SfE-Kontext auseinandergesetzt. Anhand vorhandener Konzepte aus dem Gesundheitsbereich wurden zwei Modelle entwickelt: Das erste Modell definiert vier Formen von Nachhaltigkeit auf individueller, kommunaler, organisati-

³⁶¹ Massey und Whitley (2019, S. 181) gehen im Detail auf die Vorzüge systemtheoretischer Ansätze ein, bspw. hinsichtlich der holistischen Perspektive und bzgl. der Stärke interdisziplinärer Forschungsteams. Zeitgleich äußern Massey et al. (2015, S. 29) auch Bedenken hinsichtlich der Umsetzung und Finanzierung von systemischen Evaluationen, da der Aufwand solcher Analysen „is often not congruent with how the donor industry funds projects“. In ihrem Beitrag geben die Autor*innen (ebd., S. 26ff.) einen Überblick zu systemtheoretischen Ansätzen mit dem Fokus auf Friedensentwicklung. Sie untersuchen Schulenkorfs S4D Framework (2012) sowie Lyras und Welty Peacheys SFDT (2011) und identifizieren Defizite bzgl. einer systemischen Betrachtung.

³⁶² Das S4D Framework zielt darauf ab, dass SfD Programme und Events über eine strategische Planung langfristige Nachhaltigkeit erzielen (vgl. Schulenkorf, 2012).

onaler und institutioneller Ebene, das zweite ermöglicht die Klassifizierung von Prozessen, die Nachhaltigkeit beeinflussen (u.a. im Hinblick auf den Kontrollgrad, der durch Geberorganisationen bei Programmen bestimmt wird).

Neben dem Mangel an (Wirkungs-)Studien auf anderen Ebenen und der Vernachlässigung des Themas der Nachhaltigkeit wird im SfE-Kontext immer auch das *Fehlen von Vergleichs- und Langzeitstudien* anberaumt, was auch das *Fehlen von ex-post Evaluationen* beinhaltet (vgl. Kap. 5.4.3, 5.4.4). Zudem herrschen Bedenken zu unkritischen Annahmen hinsichtlich des universellen Methoden zur Evaluation von Wirkungen (vgl., Kay, 2009, S. 1181) und „[...] to re-use social impact evaluation from Western social impact schemes, with minimal tinkering for use in low-income regions“ (Levermore, 2011, S. 351).

Dass vergleichende sowie zeitlich und inhaltlich umfangreichere Studien, v.a. mit dem Fokus auf Wirkungen und Nachhaltigkeit fehlen, liegt letztendlich an der *fehlenden Finanzierung*. Neben begrenzten Ressourcen identifizieren Whitley et al. (2019, S. 190) noch weitere Umsetzungsbarrieren, die z.T. im Folgekapitel zu SfE-Forschenden angesprochen oder bereits thematisiert wurden: herausfordernde Verhältnisse zwischen politischen Akteur*innen und beteiligten Organisationen (z.B. Angst vor negativen Ergebnissen, mangelndes Vertrauen); ungleiche Machtverhältnisse (z.B. asymmetrische oder ausbeuterische Beziehungen, neokoloniale oder neoliberale Agenden); unterschiedliche Forschungs- und Evaluationsziele (z. B. Nachweis der Programmeffizienz vs. kritische Bewertung der Programmwirkungen); longitudinale Herausforderungen (z. B. Kurzzeitprojekte, hohe Fluktuation von Teilnehmenden und Mitarbeitenden, kurzfristige Finanzierungsströme).

7.1.6 Qualität: zu wenig Meta-Evaluationen, nicht-genutzte Standards

Wie im Kapitel zur Qualität von Evaluationen dargelegt (s. Kap. 3.1.4), bieten *Evaluationsstandards* Leitlinien für eine gute Evaluationspraxis und machen damit die Qualität von Evaluationen beurteilbar (DeGEval, 2016). Als Teil der Evaluationsstandards evaluieren *Meta-Evaluationen* Evaluationen und überprüfen dadurch deren Qualität. Die Abhandlung zur EZ-Evaluation zeigt auf, dass Meta-Evaluationen im Bereich der EZ zwar

zugenommen haben, sie allerdings nach wie vor noch einen geringen Stellenwert einnehmen und Defizite in der Systematik und Transparenz aufweisen, weswegen die „[...] Evaluations-Standards erneut in den Fokus der Diskussion gerückt [...]“ und „[...] die Notwendigkeit der Umsetzung von Meta-Evaluationen stärker betont [...]“ werden sollte (Caspari, 2015, 161). Es ist nicht bekannt, in wie weit Evaluationsstandards allgemein i.R. beauftragter Evaluationen genutzt werden.

Im SfE-Kontext sind *keine expliziten Meta-Evaluationen* bekannt. Im Rahmen zweier systematischer Reviews wurde auch die methodische Qualität der Studien überprüft (vgl. Darnell et al., 2019, S. 5f.). Letztendlich sind SfE-Forschende trotz fehlender Meta-Evaluationen der Ansicht, dass sich die Qualität von Studien im Laufe der Zeit zwar verbessert hat, jedoch zahlreiche Studien immer noch Defizite oder Einschränkungen in ihrer Qualität aufweisen. Zwar existieren im SfE-Kontext Qualitätsstandards für Organisationen, die beschreiben wie M&E-Prozesse geplant, implementiert und gestaltet werden sollten (vgl. inFocus, 2014), allerdings keine, die sich an Forschende oder Evaluierende richten. Die Standards, die u.a. auf Konzepten des Nonprofit-Sektors basieren, weisen nur wenige Sportbezüge auf, was die Notwendigkeit von SfE-spezifischen Evaluationsstandards generell in Frage stellt. Auch im SfE-Kontext ist nicht bekannt, ob und inwiefern Evaluationsstandards genutzt werden.

Im Rahmen der wissenschaftlichen Begleitung wurde und wird stets versucht, wissenschaftlichen Ansprüchen gerecht zu werden, was die Ausführungen zu den *Genauigkeitsstandards* in Kapitel 6.5 veranschaulichen. Die Schilderungen zu den *Durchführbarkeitsstandards* und bzgl. der Herausforderungen i.R. der Wirkungsevaluationen zeigen jedoch Störfaktoren auf, warum diesen Ansprüchen nicht immer gerecht werden konnte. Bei Evaluationen, die bspw. über komplexere, gemischtmethodische Designs versuchen, Zuordnungslücken zu schließen, stand der Aufwand für die Beteiligten und Betroffenen aus Sicht der Autorin häufig in keinem adäquaten Verhältnis zum intendierten Nutzen der Evaluation. Eine umfassende Meta-Evaluation zu den Evaluationen der wissenschaftlichen Begleitung konnte aus vertraglichen Gegebenheiten sowie aus Gründen der persönlichen Befangenheit in dieser Arbeit nicht geleistet werden.

Aus den aufgeführten Gründen und um Evaluationsergebnisse umfangreich reflektieren und diskutieren zu können, ist es schlussendlich umso wichtiger, den*die Auftraggebenden und beteiligte Stakeholder *für wissenschaftliche Standards zu sensibilisieren* und diesbezüglich in einem fortwährenden Austausch zu bleiben. Zudem sollte gewährleistet sein, dass Evaluierende Evaluationsprozesse und -designs den praktischen Bedingungen und Beschränkungen vor Ort durch den engen Austausch mit beteiligten Stakeholdern immer flexibel anpassen können.

7.1.7 Zur Unabhängigkeit von beauftragten, externen Evaluierenden

Auch wenn interne Evaluationsabteilungen bei EZ-Agenturen seit Anfang der 1980er zur Normalität gehören (vgl. Cracknell, 2002; Stokke, 1991), wurden bereits früh externe Gutachtende rekrutiert, die Projekte berieten und evaluierten, auch im Hinblick auf der Messung von Wirkungen (Nickolmann, 2006; Reuber, 2006). Wie im Kapitel zur *Konsolidierung und Institutionalisierung* (Kap. 4.4) dargestellt, führte die zunehmende Kritik an der mangelhaften Unabhängigkeit bei Evaluationen ab den 2000ern zur vermehrten Einrichtung interner unabhängiger Abteilungen in EZ-Agenturen, wie bspw. der *Stabsstelle für Evaluierung* in der GTZ. Die Stabsstelle beauftragt seit 2005 auch Fremdevaluierungen, die durch externe Beratende bzw. Gutachtende durchgeführt werden (ebd.). Neben diesen internen unabhängigen Abteilungen und extern beauftragten Gutachtenden gründeten sich unabhängige Evaluierungsinstitute, so wie das *Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval)* im Jahr 2012. Mit der Frage, wie ein Institut unabhängig sein kann, wenn es gänzlich über den Bundehaushalt finanziert wird, war die DEval von Beginn an konfrontiert und bezog früh Stellung.³⁶³ Zudem wurde der Aspekt der *Unabhängigkeit* in den Evaluationsstandards der DEval verankert (vgl. DEval, 2018). Der Einsatz von externen Evaluierenden als *unabhängige Gutachtende* ist in der EZ in den vergangenen Jahren stetig angestiegen (vgl. Caspari, 2004, S. IX; Delhaes et al., 2017, 26. Mai). Das BMZ arbeitet mit einer Vielzahl an Gutachtenden, darunter Wissenschaftler*innen aus Universitäten und Forschungseinrichtungen, wofür die GIZ eine separate Datenbank pflegt (ebd., GIZ Stabsstelle Evaluierung, 2018b; GIZ, o.D.b). Das sog.

³⁶³ Roxin und Asche (2013) erläutern das Dilemma anhand zweier Dimensionen: Systemnähe und wissenschaftliche Unabhängigkeit. Zudem werden im Beitrag Ziele, Aufgabenbereiche sowie das Evaluierungsverständnis im Detail beschrieben.

Gutachtenmodell, wobei (Netto-)Wirkungen eines Programms seitens *externer Gutachtenden* bewertet werden, indem in der Regel Mitarbeitende von Durchführungsorganisationen oder Betroffene einmalig befragt werden, ist nach wie vor ein kostengünstiges und beliebtes Evaluationsmodell in der EZ, allerdings hinsichtlich seiner Güte, Zuverlässigkeit und Genauigkeit stark zu hinterfragen. Im Kontext der Evaluationsforschung gilt es als ein kostengünstiges, aber unzuverlässiges und ungenauestes Design (vgl. Stockmann, 2006, S. 229; 237f.; Reade, 2008, S. 15). Weniger pauschalisierend verweist Caspari (2004) auf „häufig anzutreffende methodische Defizite bei einigen EZ-Gutachter[n]“, die sie u.a. auf das *komplexe Anforderungsprofil* zurückführt: Zu den zahlreichen Rollen und Handlungsanforderungen, mit denen Evaluierende ohnehin bereits konfrontiert sind (vgl. Kap. 3.3.3), bringt die EZ-Gutachtertätigkeit zudem „einige sehr spezifische Probleme mit sich“, die den Evaluierenden noch weitere Fähigkeiten abverlangt, insbesondere im Bereich der Sozial- und Selbstkompetenz (vgl. Caspari, 2004, S. 114).

Auch im SfE-Kontext arbeiten zahlreiche SfE-Forschende als externe Evaluierende im Auftrag staatlicher EZ-Organisationen. Die Beteiligung nationaler und internationaler EZ-Organisationen beeinflusst stark, was, wie und wann im SfE-Kontext evaluiert wurde: „Practices mushroomed and donors from the Global North dominated initiatives in finding platforms for promoting their work and research agendas“ (Burnett, 2015, S. 386). Durch Initiativen der UN, bspw. i.R. des *Internationalen Jahres des Sports und Leibeserziehung* im Jahr 2005, bekamen Wissenschaftler*innen erstmalig die Möglichkeit, sich einzubringen und, wie im Falle von *Fred Coalter*, wichtige M&E-Standards zu prägen. Im weiteren Verlauf sind SfE-Forschende federführend bei Veröffentlichungen und Leitfäden des UNOSDP und später des Commonwealth Secretariat, bspw. *Bruce Kidd* und sein Forscherteam der *University of Toronto* (u.a. Simon Darnell, Peter Donnelly) (SDP IWG, 2007), *Iain Lindsey* (Lindsey & Chapman, 2017; Commonwealth Secretariat, 2019) und *Tess Kay* (Kay & Dudfield, 2013). Die Durchführung von Studien und die anschließende Veröffentlichung prägender wissenschaftlicher Artikel sind u.a. durch die Finanzierung staatlicher EZ-Agenturen möglich, wie bspw. Arbeiten von *Fred Coalter* unterstützt von *UK Sport* (z.B. Coalter & Taylor, 2010; Colater, 2017), von *Cora Burnett* finanziert durch *Sport and Recreation South Africa* und der *GTZ* (z.B. Burnett, 2009; 2013a; 2013b) und

jüngst von *Meredith Whitley* finanziert durch das *Commonwealth Secretariat* (z.B. Whitley et al., 2019a, 2020). Es stellt sich die Frage, wie das Forschungsfeld ohne den starken Einfluss der internationalen und nationalen EZ-Organisationen aussehen würde. Möglicherweise würden Schwerpunkte u.a. vermehrt auf der *Grundlagenforschung*, auf ‚*politisch-kritischeren*‘ Fragestellungen, auf der Nutzung *partizipativer Designs und Methoden*, auf der Durchführung von *Langzeitstudien* und auf *umfangreicheren Kollaborationen* liegen, wie es bereits allgemein gefordert wird (vgl. u.a. Darnell et al., 2018a; Whitley et al., 2019a).

Wie im theoretisch-konzeptionellen Kapitel zur Rolle von Evaluierenden erwähnt (s. Kap. 3.3.3), können Evaluierende als externe, ‚*unabhängige Auftragsforschende*‘ *Beeinflussungsprozessen* ausgesetzt sein, die aus der Beziehung zum Auftraggebenden resultieren: Bspw. werden gewünschte und unerwünschte Ergebnisse seitens des Auftraggebenden bereits vorab formuliert; Evaluationsinteressen bestimmter Stakeholder werden nicht oder wenig beachtet; Evaluierende werden unter Druck gesetzt, unerwünschte Ergebnisse anders darzustellen, unliebsame Befunde zu unterdrücken bzw. zu ignorieren oder Informanten zu benennen. Solche *ethischen Problem- oder Konfliktlagen* können in der Praxis häufiger vorkommen. Laut den Ergebnissen einer Befragung deutscher Evaluierenden gehört „[...] insbesondere in der EZ die Einflussnahme durch die Auftraggeber zum Alltag“ (Stockmann et al., 2011, S. 63). Auch im SfE-Kontext spricht Digel (1978, S. 15) bereits früh in einer Stellungnahme „*Erwartungshaltungen*“ seitens der Auftraggebenden an, die „mit den Möglichkeiten und Maßstäben der Sportwissenschaft nicht verträglich“ seien: „Durch ein wissenschaftliches Votum wird lediglich versucht, die politische Entscheidung symbolisch zu objektivieren“. Auch Coalter (2013a, S. 45) spielt auf die Erwartungshaltung „*of producing proof for their beliefs*“ an, die beauftragte Forschende beeinflussen und unter Druck setzen kann.

Regelwerke wie die in Kapitel 3.1.4 aufgeführten Standards für Evaluation der DeGEval (2016) beinhalten u.a. *Fairness-Standards*, die die Interaktion zwischen Evaluierenden und Stakeholdern regeln und für Auftragnehmende eine Orientierung bieten können (Stockmann, 2006; Brandt, 2007; vgl. auch Fiege, 2012, S. 240f.; Bortz & Döring, 2016, S. 976). Allerdings wird in den DeGEval-Standards die *Unabhängigkeit von Evaluierenden*,

anders als in den internationalen Standards der OECD/DAC (vgl. OECD, 2010a), weniger explizit erwähnt. Erst die oben erwähnten neuen Standards der DEval (2018) orientieren sich international und erweitern ihren Fairness-Standard um die Aspekte *Unabhängigkeit und Integrität*: „In der Evaluierungsarbeit ist das Prinzip der Unabhängigkeit ein hohes Gut, welches uneingeschränkt gilt“ (ebd., S. 2). Aufgrund dieser Tatsache sollten sich EZ-Evaluierende im deutschen Kontext an den DEval-Standards ausrichten, d.h. anders als die GIZ, die sich bislang (noch) an den DeGEval und OECD-Standards orientiert. Wichtig anzumerken ist zudem, dass es sich bei den DEval-Standards, anders als bei den als Maximalansprüche bezeichneten DeGEval-Standards (vgl. Widmer, 2000), um *anspruchsvolle Mindeststandards* handelt, die grundsätzlich für alle Evaluierungen der DEval gelten.

Sich der Vorteile als beauftragte *externe Evaluierende* bewusst (bspw. Neutralität, hohe Glaubwürdigkeit der Ergebnisse), war und ist das Evaluationsteam der DSHS i.R. seiner Arbeit jedoch auch mit Nachteilen basierend auf ihrer Rolle konfrontiert (bspw. geringe Sachkenntnisse über einzelne Projektkontexte, Abwehrreaktionen aufgrund einer wahrgenommenen ‚Kontrollinstanz von außen‘, Umsetzungsprobleme im Prozess), sowohl i.R. der Evaluationsprozesse als auch hinsichtlich der Beziehung zum Auftraggebenden. Evaluationsstandards, die u.a. auf den *Vorteil der Unabhängigkeit* von externen Evaluierenden eingehen, und bei der Rollenfindung sowie bzgl. der komplexen Handlungsanforderungen und damit verbundenen Aushandlungsprozessen unterstützen, wurden bislang nicht genutzt oder i.R. von Strategien verfolgt (vgl. Heinze et al., 2019).

7.1.8 Zur Partizipation von (zu) vielen Beteiligten

Wie im Kapitel zu *Akteur*innen im Evaluationsprozess* dargelegt (s. Kap. 3.3), sollen Evaluierende in der Lage sein, mit verschiedenen Stakeholdern und zwangsläufig unterschiedlichen Interessengruppen zu kommunizieren und ggf. zwischen ihnen zu vermitteln. Im besten Falle werden alle Beteiligte so früh wie möglich, kontinuierlich und aktiv in Evaluationsprozesse eingebunden, um eine hohe *Transparenz* und breite *Akzeptanz* zu schaffen. Hierfür werden stakeholderorientierte oder partizipative Evaluationsansätze empfohlen.

Wie in Kapitel 4.3 erläutert, hatten jene Ansätze in der allgemeinen EZ ihren Aufschwung in den 1990ern Jahren. Im weiteren Verlauf wurden zahlreiche methodische Vorgehensweisen entwickelt. Doch auch wenn die verstärkte Einbindung von Partnern zentraler Punkt der Paris Deklaration und „leave no one behind“ ein inklusives Prinzip der Agenda 2030 sind, bemühen sich nur wenige EZ-Agenturen während der Planungsphase von Evaluationen, andere Stakeholdergruppen in den weiteren Prozess einzubeziehen (vgl. OECD, 2016, S. 65; Meyer, 2020). Laut Caspari (2004, S. 114) bringen EZ-Gutachtertätigkeiten „einige sehr spezifische Probleme mit sich“, in der EZ-Tätigkeit käme es oft zu *Loyalitätskonflikten* bzgl. der beteiligten Akteur*innen.³⁶⁴

Im SfE-Kontext (vgl. Kap. 5.2) werden in diesem Zusammenhang häufig *ungleiche Machtverhältnisse* zwischen Stakeholdern thematisiert, die dazu führen, dass Geldgebende und beauftragte Evaluierende Evaluationsprozesse vorgeben, ohne weitere Beteiligte einzubeziehen, diese ausnutzen und lokale Organisationen häufig lediglich als Zugang zur Zielgruppe und ‚Datensammelnde‘ eingesetzt werden (vgl. Collison et al., 2016; Darnell et al., 2018a; Jeanes & Lindsey, 2014; Kay, 2012). Lücken zwischen praktizierenden NGOs vor Ort und Evaluierenden könnten vor allem durch die vermehrte Nutzung partizipativer Evaluationsansätze und -methoden, den Aufbau von vertrauensvollen Beziehungen und die Schaffung eines gemeinsamen Verständnisses und gemeinsamer Zielsetzungen geschlossen werden (vgl. u.a. Darnell et al., 2016; Welty Peachey & Cohen, 2016). Diese Vorstellungen bedürfen v.a. *Zeit*, die i.R. von extern finanzierten SfE-Projekten mit begrenzten Laufzeiten und anderen Ressourcen häufig knapp bemessen ist. Wie in Kapitel 5.4.4 veranschaulicht, plädieren zwar zahlreiche SfE-Forschende für *partizipative oder stakeholderorientierte Ansätze* (u.a. Collison & Marchesseault, 2018; Burnett, 2015; Darnell et al., 2016; 2018; Edwards, 2015; Kay, 2009; 2012), allerdings weisen zahlreiche, als partizipativ deklarierte Studien Defizite auf (Spaaij et al., 2018).

Bereits gesammelte, *positive Erfahrungen mit partizipativen oder stakeholderorientierten Evaluationsansätzen* konnten in der Praxis der wissenschaftlichen Begleitung mal mehr, mal weniger umgesetzt werden. Die partizipativen Prozesse waren häufig mit CD-

³⁶⁴ In den *Ethischen Leitlinien der Arbeitsgemeinschaft Entwicklungsethnologie (AGEE)* werden ethische Dilemmata in der EZ-Praxis angesprochen und Handlungsziele formuliert werden (vgl. AGEE, 2013).

Maßnahmen zu M&E verbunden, bei denen beteiligte Stakeholder sensibilisiert („Evaluation ist nicht nur Kontrolle!“)³⁶⁵ und geschult sowie Designs und Instrumente gemeinsam entwickelt werden konnten. Wie oben bereits verdeutlicht, benötigen diese Prozesse *Zeit und weitere Ressourcen* (Geld, Personal), die nicht in allen Projekten zur Verfügung standen. Die Erfahrungen verdeutlichen, dass die Beteiligung einer sinnigen Auswahl von Akteur*innen von großem Mehrwert sein kann, wenn Zuständigkeiten und Aufgaben klar festgelegt sind. Sind zu viele Akteur*innen ohne klare Rollen und Aufgaben an Evaluationsprozessen beteiligt, kann dies negative Auswirkungen auf den Prozess und die Qualität der Evaluation haben.

Letztendlich müssen sich SfE-Forschende oder Evaluierende zum einen mit der Tatsache auseinandersetzen, dass sie es *nie* allen beteiligten Stakeholdergruppen ‚recht machen‘ können. Ihre Gründe, im SfE-Kontext zu forschen, stimmen vermutlich in den meisten Fällen nicht mit jenen derer überein, die Evaluationen beauftragen. Hier sind Wissenschaftler*innen an die ToR der Auftraggebenden gebunden, i.R. derer allerdings auf Evaluationsstandards verwiesen werden kann, die bspw. die Partizipation von Stakeholdern in den Vordergrund rückt. Des Weiteren sind die Gründe möglicherweise andere als die Motivationen, Zwänge und Interessen von lokalen Stakeholdern (vgl. Collison et al., 2016). Neuralgischer Punkt in der Debatte darum, welchen Stakeholdern SfE-Forschende und Evaluierende gerecht werden, ist und bleibt, mit allen Beteiligten transparent zu klären und zu formulieren, warum und in welchem Interesse Forschungs- oder Evaluationsaktivitäten durchgeführt werden: „[...] it is important to ask what, why and for whom research is important or even needed?“ (Darnell et al., 2018a, S. 120).

Hilfreich in diesem Zusammenhang ist es, wenn Forschende eine *größere Reflexivität* entwickeln, was die eigene Rolle als Forschende an sich und die Auswirkungen auf die durchgeführte Forschung betrifft. Diese findet als wichtiger Aspekt bei der Entwicklung und Durchführung von Studien im SfE-Kontext bislang zu wenig Beachtung und wird zu wenig diskutiert (Darnell et al., 2018b). Ein „reflexive sense of humility“ (Kidd, 2011, S. 609) würde u.a. für transparente(re) Evaluationsergebnisse sorgen, einschließlich der

³⁶⁵ Dadurch, dass die Evaluationsaktivitäten und -funktionen besprochen wurden, stießen die M&E-Aktivitäten insgesamt auf eine breitere Akzeptanz bei den Stakeholdern und die zuvor alleinig assoziierte Kontrollfunktion bzgl. der Evaluationen erweiterte sich auf weitere.

Berichterstattung von negativen Ergebnissen (Whitley et al., 2019b),³⁶⁶ und zu einer zunehmenden kritischen Auseinandersetzung mit dem Forschungsfeld, wie sie bislang der Fall ist (Darnell et al., 2018b).

7.2 Schlussfolgerungen und Ausblicke

Das Forschungsfeld SfE ist auf der einen Seite durch den starken (finanziellen) Einfluss internationaler und nationaler EZ Agenturen geprägt, die den Sport zur Erreichung von Entwicklungszielen i.R. ihrer Maßnahmen einsetzen. Evaluationsforschung wird hierbei im Sinne ihrer *Anwendungsorientierung* genutzt, wobei Evaluationen hauptsächlich eine *Legitimationsfunktion* einnehmen. Auf der anderen Seite ist das Forschungsfeld seitens der Forschenden einem ‚*konfusen*‘ *Verwissenschaftlichungsprozess* ausgesetzt, da Forschende jeglicher Disziplinen den Bereich SfE als eine Art ‚Spielwiese‘ nutzen und versuchen, ihre theoretischen Ansätze und Methoden zu platzieren. Es bleibt die Frage, ob es dem Forschungsfeld gelingen wird, sich aus diesem Spannungsfeld zu lösen und sich weiterzuentwickeln. Hierzu wären Veränderungen auf beiden Seiten notwendig, d.h. seitens der Forschenden und seitens der EZ- Akteur*innen:

Von *Seiten der Wissenschaft* bleibt die Tatsache bestehen, dass das Forschungsfeld weiterhin von den verschiedensten Teildisziplinen bearbeitet werden wird, da es sich mit völlig unterschiedlichen Forschungsobjekten (z.B. Maßnahmen, Zielgruppen, Organisationen, Policies) und Phänomenen auf unterschiedlichen Ebenen (mikro, meso, makro) beschäftigt. Weiterentwicklung ist möglich, wenn es dem Forschungsfeld gelingt, sich von *Multidisziplinarität* hin zu einer *Interdisziplinarität*³⁶⁷ zu entwickeln. D.h., dass künftig vermehrt in *Forschungsteams* gearbeitet werden sollte (vgl. Burnett, 2015, S. 388; Massey & Whitley, 2019, S. 181; Whitley et al., 2019b, S. 11) und dass wissenschaftliche Fragestellungen im Kontext SfE nicht mehr nur unabhängig voneinander und ohne methodischen, terminologischen und konzeptionellen Austausch stattfindet, sondern wie

³⁶⁶ Das Berichten von negativen Ergebnissen wird u.a. als risikoreich angesehen, da von Evaluationsergebnissen für NGOs bspw. eine Weiterfinanzierung abhängt (Welty Peachey & Cohen, 2016).

³⁶⁷ *Interdisziplinarität* definiert sich nach Defila, DiGiulio und Scheuermann (2008, S. 12) als ein „[...] integrationsorientiertes Zusammenwirken von Personen aus mindestens zwei Disziplinen im Hinblick auf gemeinsame Ziele und Ereignisse, in dem die disziplinären Sichtweisen zu einer Gesamtsicht zusammengeführt werden“. „Darüber hinaus wird von einer interdisziplinären Forschungssituation gesprochen, wenn das Problem sowie die Methode einer anderen Theorie entspringt“ (Parthey, 2011 zitiert in Barisch-Fritz & Volk, 2016, S. 16).

in den Sportwissenschaften auch Fragestellungen und komplexe Problemstellungen interdisziplinär bearbeitet werden. Wichtig hierfür sind folgende Schritte (vgl. Barisch-Fritz & Volk, 2016, S. 23):

(1) Kommunikation mit dem Ziel, eine gemeinsame Sprache zu finden

Hierzu gehört ein Verständnis für fremdes Fachwissen (bspw. zu Grundlagenforschung) und die Kompetenz, das eigene Fachwissen gut vermitteln zu können (bspw. die Evaluationsforschung). Wie die Ausführungen zu den Definitionen verdeutlichen, basieren viele Diskussionen und Auseinandersetzungen bzgl. des Forschungsfeldes auf *definitorischen Ungenauigkeiten*. D.h. es muss klarer differenziert und kommuniziert werden, was gemacht wird (Grundlagenforschung oder Evaluationsforschung), nur so kann ein gegenseitiges Verständnis entstehen. Genauso wichtig ist es, dass SfE-Forschende und EZ-Akteur*innen eine gemeinsame Sprache sprechen: Wenn EZ-Akteur*innen – egal, ob staatliche Agenturen oder nichtstaatliche NGOs – mit SfE-Forschenden/Evaluierenden zusammenarbeiten, muss vorab geklärt werden, was bspw. unter ‚Wirkungen‘, ‚direkte/indirekte Zielgruppen‘ und möglicherweise unter ‚SfE‘ selbst verstanden wird. Des Weiteren sollte transparent zwischen Evaluierenden und Auftraggebenden *Funktion bzw. Nutzen* von Evaluationen kommuniziert und möglicherweise dokumentiert werden. Evaluierenden wäre dann möglicherweise die Entscheidung gegeben, ob sie bspw. Evaluationen lediglich zu Legitimationszwecken durchführen wollen, d.h. auch mit der Erwartung verbunden, ‚positive‘ Ergebnisse zu generieren, oder ob sie von dieser Aufgabe aufgrund eigener Rollenvorstellungen zurücktreten.

(2) ‚Blick nach rechts und links‘ mit dem Ziel, theoretische Ansätze und Methoden zu vereinheitlichen

Das ‚Eindenken‘ in und das Verstehen von theoretisch-konzeptionellen Vorgehensweisen anderer Disziplinen bedarf der Bereitschaft und Zeit hierfür. Statt weiterhin nur Bekanntes aus der eigenen Disziplin zu nutzen und dies einem anderen Forschungsfeld ‚überzustülpen‘ zu wollen, sollten Forschende sich gerade aufgrund der Nähe zu EZ-Institutionen und der damit einhergehenden geteilten Ausrichtung an den Nachhaltigen Entwicklungszielen tiefer mit bewährten Modellen aus der EZ und/oder modifizierten

theoretischen Modelle in Form von SfE-Schlüsselkonzepten beschäftigen und diese vermehrt anwenden. Vor allem, wenn im SfE-Kontext *Evaluationsforschung* betrieben wird, sollten EZ-Kenntnisse (inklusive Definitionen, Standards, Modelle) mehr genutzt werden. Für den Bereich der *Grundlagenforschung* ist allerdings bislang noch unklar, welche Ansätze und Methoden sich wirklich bewähren, da hierzu die Positionen stark variieren. Als erster Schritt könnten die als ‚Schlüsselkonzepte‘ bezeichneten Ansätze (vgl. Welty Peachey et al., 2019) genauer untersucht und eingeordnet werden, auch im Hinblick auf ihre Umsetzbarkeit i.R. anderer Studien. Da sich das Feld SfE auch nach wie vor im Hinblick auf seine Wirksamkeit und seinen Beitrag zu entwicklungspolitischen Zielsetzungen ‚beweisen‘ muss, gilt es zum einen ausgewiesene *Wirkungsevaluationen zu evaluieren (Metaevaluation)*. Fragen, wie ‚Welche Ansätze und Methoden wurden in welchen Kontexten und auf welchen Ebenen genutzt?‘, ‚Was sind Vor- und Nachteile?‘ könnten so beantwortet werden. Des Weiteren gilt es, einen genaueren Blick darauf zu werfen, welche bereits *bestehenden Modelle im EZ-Kontext* genutzt werden und wurden, um Wirkungen zu messen (s. Oberndörfer et al., 2010). Da neben allgemeinen Evaluationsforschenden auch SfE-Forschende in *systemtheoretischen Ansätzen* Lösungen bzgl. des Mikro-Makro-Paradoxons sehen, sollten auch diese als Grundlage für künftige Forschungsaktivitäten in Erwägung gezogen werden. Zuletzt sollten auch im SfE-Kontext *mehr ex-post Evaluationen* i.R. abgeschlossener Projekte durchgeführt werden, um *Prozesse* (Welchen Einfluss haben Aspekte in der Planung, Steuerung und Nachsorge der Projekte auf ihre Nachhaltigkeit?), *Nachhaltigkeit* (Was passiert mit den Projekten, den SfE-Ansätzen und Methoden nach Ende der GIZ-Finanzierung?) und ihre *Gesamtwirkung* (Welche intendierten und nicht-intendierten Wirkungen können festgestellt werden?) bewerten zu können.

(3) *Gegenstandsbeschreibung mit dem Ziel, gemeinsame Problemsichten zu finden*

Hierfür muss die unterschiedliche disziplinspezifische Strukturierung der Realität erkannt werden. Möglicherweise bildet der Beitrag von Sport zu ausgewählten Nachhaltigen Entwicklungszielen und die weitreichende Ausrichtung von SfE-Organisationen an Themen wie sozialer Ungleichheit, Umwelt, safeguarding, Geflüchtete und sozialem Unternehmertum (vgl. Giulianotti et al., 2019) gemeinsame Nenner. Gerade die *SDGs als weltumspannende, ebenen- (mikro, meso, makro), sektoren- und disziplinübergreifende*

Zielsetzungen bieten die Möglichkeit, gemeinsame Problemsichten zu finden. Deshalb gilt es diesbezüglich, vor allem im Hinblick auf die Weiterentwicklung und letztendliche Vereinheitlichung von theoretischen Ansätzen und Methoden, sich an neuen Konzepten der Evaluationsforschung, der Entwicklungs- und Nachhaltigkeitsforschung zu orientieren (vgl. Meyer, 2020).

Für die Seite der *EZ-Akteur*innen* können an dieser Stelle lediglich *wünschenswerte* Entwicklungen formuliert werden. Sicher ist jedoch, dass EZ-Organisationen ‚den Sport‘ voraussichtlich auch in den kommenden Jahren als ein niederschwelliges und kosteneffizientes Instrument i.R. von Projekten und Programmen zur Erreichung entwicklungspolitischer Zielsetzungen einsetzen werden. Dementsprechend wird (Evaluations-)Forschung im Kontext SfE auch in Zukunft maßgeblich von Geldern staatlicher EZ-Organisationen (mit)finanziert und anwendungsorientiert geprägt sein. Dies birgt die Gefahr, dass Evaluationen weiterhin hauptsächlich taktisch-strategisch instrumentalisiert sind und vor allem im Sinne der Legitimation bzw. Rechenschaftslegung bzgl. Steuergelder und der ‚Erfolgskontrolle‘ genutzt werden – und weniger im Sinne ihrer Lern- und Dialog- oder Erkenntnisfunktion. Vor diesem Hintergrund wird das Forschungsfeld SfE durch die *finanzielle Abhängigkeit* dieser Geldgebenden kaum die Möglichkeit haben, sich weiterzuentwickeln. Es sei denn, EZ-Akteur*innen *rücken von einer reinen Anwendungsorientierung ab* und *öffnen sich mehr für eine Verwissenschaftlichung* der Prozesse. D.h., dass bspw. mehr finanzielle Mittel für die Grundlagen- und Langzeitforschung zur Verfügung gestellt werden und dass man der Beratung bezüglich Fragestellungen, Designs und methodischer Vorgehensweisen offen begegnet, statt sich auf starre Zielsysteme zu fixieren, die Methoden, Fragestellungen etc. vorgeben. Hoffnung gibt, dass vor dem Hintergrund der Agenda 2030 neben den steuernden und legitimierenden Evaluationsfunktionen auch die erkenntnisbringende *Funktion zur Aufklärung* von Gesellschaften zunehmend an Bedeutung gewinnen wird, bei der die Generierung von objektivem Wissen im Vordergrund steht und sie dadurch die größte Schnittstelle zur Grundlagenforschung aufweist. Meyer und Stockmann (2017, S. 254) prognostizieren, dass Evaluation künftig stärker als „Instrument sozialer Reflexion“ eingesetzt werden wird, wofür allerdings erst Räume zur Diskussion von sozialen Problemen und politischen Lösungen geschaffen werden müssen (vgl. ebd., S. 244). Diese Aufklärung durch Evaluation kann jedoch nur

stattfinden, „[...] wenn Evaluation *von der Zwangsjacke der Auftraggeberwünsche befreit* ist“ (Stockmann, 2017, S. 44).³⁶⁸

Da gerade Wirkungsevaluationen aufgrund der Legitimationsfunktion weiterhin gefragt sein werden, sollte eine *Offenheit bzgl. der Machbarkeit und Durchführbarkeit* i.R. von Beratungen auf Seiten der Auftraggebenden gegeben sein, bspw. bzgl. der Zielgruppe oder wenn experimentelle Designs nicht immer Goldstandard sein können, v.a., wenn man in pädagogischen Settings neben praktischen und methodischen auch mit ethischen Herausforderungen konfrontiert ist. Für beauftragte Evaluierende bleibt einzig die strategische Möglichkeit, *Aushandlungs- und (Auf-)Klärungsformate* zu schaffen, i.R. derer bspw. Erwartungen, Ansprüche, Handlungsanforderungen, Evaluationszwecke oder -fragestellungen regelmäßig diskutiert werden können (vgl. Heinze et al., 2019).

Selbstverständlich werden auch *Entwicklungen im Hochschulkontext* einen großen Einfluss darauf haben, inwiefern sich das Forschungsfeld SfE künftig im Spannungsfeld zwischen Verwissenschaftlichung und Anwendungsorientierung weiterentwickeln und positionieren wird. Bislang sehen (Nachwuchs-)Wissenschaftler*innen das Hochschulsystem selbst, in dem SfE als Forschungsfeld oft nicht anerkannt ist, Forschungsgelder fehlen und Auftragsforschungen kaum Publikationsmöglichkeiten bieten, als größte Hürde an, sich als SfE-Forschende weiter zu engagieren (vgl. Welty Peachey & Cohen, 2015). Eine Erörterung zum Finanzierungssystem von Hochschulen und zur Ressourcenverteilung i.R. verschiedener Forschungsfelder würde den Rahmen der Arbeit übersteigen. Angemerkt werden soll jedoch, dass die institutionelle Situation an der DSHS als Europas größte Sportuniversität zwar eine einzigartige ist, dass das Forschungsfeld SfE aber auch hier in den vergangenen Jahren hauptsächlich durch Drittmittelfinanzierungen ausgebaut werden konnte. Aufgrund dieser Abhängigkeit ist es möglich, dass sich das Forschungsfeld an Hochschulen, auch an der DSHS, sehr langsam weiterentwickeln wird. Zudem ist wahrscheinlich, dass es aufgrund seiner Nähe zu unterschiedlichen Disziplinen weiterhin als ‚Nebenzweig‘ bearbeitet wird, bspw. im Bereich der Sportsoziologie, des Sportmanagements oder der Sportpädagogik, was wiederum stark an das *Forschungsinteresse einzelner Personen* gekoppelt ist. Auch im Hinblick auf der Entwicklung

³⁶⁸ Hervorhebungen durch die Autorin.

des Forschungsfeldes an Hochschulen steckt *Hoffnung in der Agenda 2030*: Hochschulen werden als Schlüsselakteur*innen für nachhaltige Entwicklung gesehen, bspw. auch im *Nationalen Aktionsplan Bildung für Nachhaltige Entwicklung* (vgl. bne-portal, o.D., DAAD, Mai 2018). Dies bedeutet, dass das Thema Nachhaltigkeit und die SDGs hier weiterhin und möglicherweise mit steigender Tendenz als Forschungsgegenstand unterschiedlicher Disziplinen Bearbeitung finden wird, wünschenswerterweise mehr als bislang durch die öffentliche Hand finanziert.

Forschende im SfE-Kontext kommen, wie oben beschrieben, aus den unterschiedlichen Disziplinen, haben sehr verschiedene Werdegänge, ihre Forschung wird meist durch EZ-Akteur*innen i.R. von Drittmittelprojekten an Universitäten finanziert oder sie arbeiten als freie Gutachtende. Die zuletzt aufgeführten Problemlagen bzgl. der Qualität von Evaluationen, der Unabhängigkeit und der Einbeziehung von Stakeholdergruppen führen zur selben Ressource: *den Evaluationsstandards*, die alle diskutierten Aspekte thematisieren. Dies bedeutet, dass Forschende und Evaluierende im SfE-Kontext sich v.a., um der *Verwissenschaftlichung Willen* sowie um die transparente und nachvollziehbare *Einbindung von Beteiligten* zu gewährleisten und die *eigene Objektivität und Unabhängigkeit* zu wahren, an *allgemeinen Evaluationsstandards* orientieren sollten, SfE-spezifische sind hierfür nicht notwendig. Im deutschen Kontext empfehlen sich die Standards der DEval (2018), im internationalen Kontext jene der OECD/DAC (OECD, 2010a), da beide einen Schwerpunkt auf die *Unabhängigkeit* von Evaluierenden legen. Die Kriterien der Standards (1. Nützlichkeit der Evaluationen; 2. Gewährleistung der Evaluierbarkeit; 3. Fairness gegenüber aller Beteiligten und die Wahrung der Unabhängigkeit und Integrität; 4. Genauigkeit, Wissenschaftlichkeit und Nachvollziehbarkeit; (5) Vergleichbarkeit) sind alle in den Diskussionsaspekten wiederzufinden und sollten demnach bei beauftragten Evaluationen von Anfang an, bspw. i.R. der ToR als Referenz, aber auch im fortlaufenden Prozess (bspw. in Form regulärer Besprechungen) einbezogen werden. Des Weiteren wäre es für SfE-Forschende bzgl. ihrer Rollenfindung in Spannungsfeldern hilfreich, wenn sie sich erstens mit den *Evaluierungspolicies der beauftragten Unternehmen* im Detail auseinandersetzen, um einen Überblick darüber zu bekommen, wie die M&E-Struktur aufgebaut ist und abläuft, und wie interne und externe Evaluierungsaufgaben und -zuständigkeiten geregelt sind. Dies dient v.a. auch dazu, dass eigene Rollen und

Aufgaben abseits der ToR genau geregelt sind. Zweitens sollten sie vor dem Hintergrund der Beteiligung unterschiedlicher Stakeholdergruppen *partizipative Prozesse* im Evaluationskontext aufgrund ihrer Mehrwerte und vor dem Hintergrund von EZ-Verpflichtungen i.R. internationaler Agenden (bspw. Paris Deklaration) ‚*pushen*‘ (bspw. durch die Einbindung in ToRs). Zeitgleich müssen sie sich bewusst sein, was für einen *Mehraufwand* (Zeit, Gespräche, Feedback) *und Verpflichtungen* dies mit sich bringt und dies klar mit allen Beteiligten kommunizieren. Und drittens und letztens, müssen SfE-Forschende auch in einem eher anwendungsorientierten Kontext stets *um die Qualität von Evaluationen bemüht* sein, d.h. sich generell mehr mit *Meta-Evaluationen* auseinandersetzen, Meta-Evaluationen durchführen oder die eigenen Evaluationen auf einer Meta-Ebene qualitativ hinterfragen.

Ein weiterer wichtiger Aspekt, der an dieser Stelle benannt werden soll, und mit dem sich SfE-Forschende bzw. Evaluierende auseinandersetzen sollten, ist die eigene *Reflexivität*, die zunehmend Eingang im allgemeinen wissenschaftlichen Kontext (vgl. Dahler-Larsen, 2017, S. 30) und auch in Studien zu SfE findet (vgl. Darnell et al., 2018b; Kidd, 2011, S. 609; Whitley et al., 2019b). „Selbiges kann Evaluatoren helfen, ihre Rolle im Wissensgenerierungsprozess einer komplexen Gesellschaft zu verstehen“ (Dahler-Larsen, 2017, S. 30). Dies geht damit einher, dass SfE-Forschende sich ihrer Identität, inklusive ihres Ausbildungshintergrunds, ihrer Disziplin, ihrer Herkunft, etc. bewusst sind und diese bzgl. ihrer unterschiedlichen Forschungsaufgaben angemessen und umfangreich i.R. ihrer Arbeit reflektieren. Sie müssen sich ihrer unterschiedlichen Rollen bewusst sein, die zwangsläufig Spannungsfelder erzeugen. In diesem Zusammenhang soll v.a. auf die *enge Beziehung zwischen Evaluation und Politik und deren Auswirkung auf Evaluierende* hingewiesen werden: Wenn in einem politischen Kontext evaluiert wird, können Evaluierende eine *value-neutral*, *value-sensitive* oder *value-critical* Position einnehmen. Im Rahmen der ersten Position werden Politik und Evaluation strikt getrennt, d.h. der Evaluierende arbeitet unabhängig, um Maßnahmen objektiv und neutral zu bewerten. Bei der zweiten Position wird angenommen, dass die Evaluation in einem politischen Umfeld stattfindet und nicht komplett von diesem getrennt werden kann, vor allem, wenn es um die Bewertung von Maßnahmen geht. Letztere begreift Evaluation und Po-

litik als untrennbar, wobei Evaluierende nicht neutral werten können und eigene ethische und moralische Einstellungen während des Evaluationsprozesses transparent formulieren sollten (Vestman & Conner, 2007). Zur Positionierung sollten sowohl eine persönliche als auch eine gemeinschaftliche Auseinandersetzung (im Evaluationsteam) stattfinden.

Schlussendlich decken sich die diskutierten Aspekte rund um SfE-Forschende mit zwei aktuellen Grundsatzkonflikten in der Evaluationsforschung: dem zwischen *Beteiligung und Unabhängigkeit* und jenem zwischen *Wissenschaft und Praxis* (Meyer & Stockmann, 2017). Bzgl. des ersteren Grundsatzkonflikts können Stockmann und Meyer (2017) aufgrund großer sektoraler Unterschiede keinen eindeutigen Trend ausmachen. Einerseits wäre es möglich, dass Organisationen sich von externen Evaluationen zugunsten interner Evaluationen abkehren. Das andere Extrem wäre, dass sich die Partizipation in Evaluationsverfahren professionalisiert. Damit der Trend eher in Richtung Professionalisierung geht, müsste an bestehenden theoretischen und methodischen Defiziten hinsichtlich der Beteiligungskonzepte gearbeitet werden (ebd.). Für den sportbezogenen EZ-Kontext ist anzunehmen, dass auch künftig mit externen SfE-Forschenden zusammengearbeitet werden wird. Eine Professionalisierung von Partizipation in Evaluationsverfahren würde in diesem Zusammenhang bedeuten, dass SfE-Forschende bereits existente Ansätze und Modelle, die Partizipation inklusive der Reflexivität seitens der Forschenden thematisieren (vgl. Schulenkorf et al., 2020; Spaaij et al., 2018), vermehrt nutzen und damit an den theoretischen und methodischen Defiziten hinsichtlich der Partizipationskonzepte gearbeitet wird. Zudem sollte die Beteiligung von Stakeholdern von Beginn an mit Auftraggebenden kommuniziert werden.

Im Hinblick auf den Grundsatzkonflikt *Wissenschaft und Praxis*, der Hauptbetrachtungsgegenstand dieser Arbeit ist, geht der Trend in der Evaluationsforschung einerseits dahin, dass mehr Wissenschaftlichkeit gefordert sein wird, d.h. komplexere und qualitativ hochwertigere Evaluationen. Dies wird zwangsläufig auch von beauftragten SfE-Forschenden verlangt werden. Andererseits geht der Trend zu einer stärkeren praktischen Verwertbarkeit, bspw., dass Evaluationsergebnisse intern zum organisationalen Lernen und extern zur Rechenschaftslegung genutzt werden können. Stockmann und Meyer

(2017, S. 175) sind der Auffassung, dass „[...] die Spannungen zwischen Wissenschaft und Praxis in der Evaluation eher zu als ab[nehmen]. [...] Ob dabei die *praxisorientierte Servicefunktion* oder die *qualitätsorientierte Wissenschaftlichkeit* die Oberhand behält, liegt wesentlich an der *Marktentwicklung* und der *Evaluationskultur* auf Seiten der Auftraggeber“ (ebd., S. 175).³⁶⁹

7.3 Reflexion und Perspektiven

Im Folgenden sollen zum einen Möglichkeiten und Grenzen der Autorin als involvierte Evaluierende reflektiert werden. Zum anderen geht es darum, selbiges bzgl. der Rahmung der Arbeit vorzunehmen, was die theoretisch-konzeptionelle sowie die methodische Vorgehensweise einschließt. Diesbezügliche Perspektiven, bspw. für weitere Forschungsarbeiten, werden im Anschluss aufgezeigt.

7.3.1 Möglichkeiten und Grenzen

Als *involvierte Evaluierende* kann festgehalten werden, dass die wissenschaftliche Begleitung des SV SfE der GIZ der Autorin Einblicke in die Entwicklung der sportbezogenen EZ in Deutschland sowie in unterschiedliche Prozesse, darunter managementorientierte, strategische und M&E-spezifische bot und bietet. Aufgrund dieser Möglichkeiten wurde die eigene Projektarbeit neben anderen Untersuchungsfoki als Gegenstand dieser Arbeit gewählt. Der Zugang ermöglicht die umfangreiche Darstellung von SfE in Deutschland (Kapitel 2) und die Beschreibung der evaluationsspezifischen Aktivitäten (Kapitel 6). Auch wenn versucht wurde, Entwicklungen, Prozesse und Aktivitäten so neutral und objektiv wie möglich darzustellen, kann eine subjektive Färbung und möglicherweise eine Selektion bei der Hervorhebung von Schwerpunkten ob der persönlichen Involvierung nicht ausgeschlossen werden. Eine umfangreiche Reflexion im Sinne der geforderten Reflexivität von Evaluierenden und Forschenden kann nicht umfangreiche stattfinden, da diese Informationen z.T. der vertraglich festgelegten *Verschwiegenheit* unterliegen (vgl. GIZ, Juli 2018).

³⁶⁹ Hervorhebungen durch die Autorin.

Die in Kapitel 3 eingeführten *Untersuchungsperspektiven* nutzen Grundaspekte der Evaluationsforschung und haben im Hinblick ihrer Systematik eine ordnende und analytische Funktion für diese Arbeit. Durch die angeführten Bestandteile der Evaluationsforschung sowie der Betrachtung der Beteiligten, die einen maßgeblichen Einfluss auf Evaluationsprozesse haben können, konnten Evaluationsaktivitäten und Akteur*innen in den unterschiedlichen Kontexten (EZ allgemein, SfE, wissenschaftliche Begleitung des SV SfE) im Detail untersucht und die zu Beginn aufgeführten Forschungsfragen beantwortet werden. Das genutzte Modell arbeitete definitorische Schwächen, der Mangel an ex-post Evaluationen (Ansatz) und Herausforderungen im Hinblick auf die Evaluation von Wirkungen und Nachhaltigkeit und damit verbundene Einflussfaktoren (Organisation/intern, Umwelt/extern) heraus. Auch Problematiken bzgl. unterschiedlicher oder einseitiger Funktionszuschreibungen sowie eingebundener Stakeholder wurden ersichtlich. Der Fokus auf die Qualitätssicherung von Evaluationen und die damit einhergehende Nutzung der Evaluationsstandards verdeutlichten Missstände im allgemeinen und sportbezogenen EZ-Kontext. Doch vor allem bei der Betrachtung des SfE-Kontexts stößt das Modell an Grenzen: Eindeutige Zuordnungen oder klare Typisierungen von Studien sind selten möglich, da hier die *Grenzen zwischen Forschung und Evaluation fließend* sind. Schnell wurde deutlich, dass der SfE-Kontext nicht nur aus der Perspektive der Evaluationsforschung betrachtet werden kann. Mit der Absicht, das Forschungsfeld SfE möglichst in seiner Gesamtheit zu fassen, stößt der*die Leser*in hier möglicherweise auf begriffliche Abweichungen. Dadurch, dass in dieser Arbeit allerdings eine *Perspektive aus der Evaluationsforschung* gewählt wurde, wurde deutlich, dass im SfE-Kontext weit mehr Evaluationsforschung betrieben wird als deklariert.

Letztendlich setzt das Modell an zwei wesentlichen Aspekten an, die anfangs als *Forschungsdefizite* beschrieben wurden (vgl. Astbury, 2017; Heinze et al., 2019; Meyer & Stockmann, 2017): Es nimmt zum einen die Evaluationsforschung an sich, die Evaluationspraxis und die jeweilig zentralen Probleme in den Blick. Des Weiteren befasst es sich mit der bislang vernachlässigten Rolle und Beeinflussung von/durch Personen, die an Evaluationsprozessen beteiligt sind, mit dem Fokus auf beauftragten Evaluierenden.

Der Genese der Promotionsarbeit geschuldet, konnte methodisch nicht so strukturiert vorgegangen werden, wie es bei theoretisch-konzeptionellen Arbeiten möglicherweise üblich ist. Anfänglich war eine empirische Untersuchung angedacht, die aufgrund äußerer Umstände jedoch nicht umgesetzt werden konnte. Deshalb wurden als theoretische Fundierung gedachte Kapitel (EZ-Evaluation, Evaluation und Forschung im SfE-Kontext) zu Hauptgegenständen der Arbeit gemacht. Wären die Überlegungen von Beginn an, an einer theoretischen, also einer rein literaturbasierten Arbeit ausgerichtet gewesen, wäre die Autorin im Nachhinein methodisch anders, mehr im Sinne einer systematischen Review, vorgegangen.

7.3.2 Theoretisch-konzeptionelle Perspektiven

Die Autorin ist der Ansicht, dass das genutzte Modell in seinen *Teilbereichen* für weitere Forschungsarbeiten mehr Perspektiven birgt als im gesamten:

Die Teile zu Definitionen, Ansätzen, Funktionen und Evaluationsgegenständen bieten eine systematische Grundlage dafür, Evaluationen in unterschiedlichen Anwendungsbereichen zu bestimmen und damit zu typisieren. Besonders wenn es um die Evaluation von Maßnahmen (Projekten, Programmen) geht, eignet sich die Programmverlaufsperspektive, also Evaluationsaktivitäten in unterschiedlichen Phasen zu betrachten, vor allem mit dem Schwerpunkt auf der ex-post Wirkungsphase.

Die Verlaufsperspektive sowie die Berücksichtigung interner und externer Einflussfaktoren auf die Wirksamkeit von Maßnahmen wurden dem *Ansatz zur Wirkungsevaluation des Centrums für Evaluation (CEval)* entnommen, der auf theoretischen Ansätzen aus der Lebensverlaufs-, Organisations- und Innovations-Diffusionsforschung sowie einem mehrdimensionalen Nachhaltigkeitsmodell beruht (Stockmann, 2006). Da der CEval-Ansatz, auch *Stockmann'scher Ansatz* genannt, bereits in einer Vielzahl von EZ-Bereichen zum Einsatz kam,³⁷⁰ könnte er perspektivisch modifiziert und auch im SfE-Kontext eingesetzt werden. Einzelne SfE-Projekte, aber auch ganze Programmlinien und -strategien

³⁷⁰ Der CEval-Ansatz wurde bereits in den EZ-(Teil-)Sektoren Bildung, Gesundheitswesen, Wasserversorgung/Abwasser, Ressourcenmanagement/Umweltschutz, Ländliche Entwicklung, Humanitäre Hilfe sowie auch wenig sektorspezifisch oder ganz sektorunspezifisch genutzt (Oberndörfer et al., 2010, S. 218).

von Organisationen könnten mit dem Ansatz auf ihre Effizienz, Wirksamkeit und Nachhaltigkeit hin untersucht und bewertet werden (vgl. Silvestrini & Reade, 2008). Auch Oberndörfer et al. (2010) betonen den flexiblen Einsatz des Ansatzes in verschiedenen Bereichen und auf unterschiedlichen Ebenen (Mikro-Meso-Makro) sowie die „klare/gute“ Begründung der Wirksamkeit (ebd., S. 215ff., S. 229).

Argumente für die Erprobung und Nutzung des CEval-Ansatzes im SfE-Kontext sind aus Sicht der Autorin in allen Teilkonzepten vorhanden: Die Vorteile bzgl. der Betrachtung von Maßnahmen in einer *Verlaufsperspektive*, wodurch kausale Verkettungen deutlich werden und die ‚vernachlässigte‘ Nachförderphase integraler Bestandteil ist, sowie die Betrachtung von *(organisations-)internen und externen Faktoren*, die einen maßgeblichen Einfluss auf die Wirksamkeit und Nachhaltigkeit von Maßnahmen haben (vgl. Stockmann, 2006, S. 107ff., 112ff.), wurden in dieser Arbeit bereits erwähnt und wahrgenommen. Des Weiteren wird ‚der Sport‘ im EZ-Kontext nach wie vor noch als Novum und damit als ‚innovatives‘ Medium betrachtet, was sich mit dem *innovations- und diffusionstheoretischen Ansatz* deckt. Sport kann in diesem Zusammenhang als „Verfahrensinnovation“ in der EZ bezeichnet werden, der aufgrund ihrer *relativen Vorteilhaftigkeit* (Sport als kostengünstiges und niederschwelliges Medium und damit die ‚bessere Alternative‘ zu anderen EZ-Instrumenten), ihrer *Vereinbarkeit/Kompatibilität* (Sport stimmt mit existierenden Werten, Erfahrungen und Bedürfnissen von Nutzer*innen überein) und ihrer *geringen Komplexität* (Sport wird als ‚universal language‘ von jedem verstanden und ist leicht auszuführen) ein hohes Diffusionspotenzial zugesprochen wird (vgl. Mohr, 1977, S. 60). Zentrales Argument ist, dass der CEval-Ansatz den im SfE-Kontext bislang *vernachlässigten Nachhaltigkeitsaspekt* umfasst. Er beinhaltet operationale Parameter bzgl. vier Dimensionen von Nachhaltigkeit, was über die unzureichende Betrachtung von Nachhaltigkeit in der EZ hinausgeht (Caspari, 2004; Stockmann, 2006) und *Überschneidungen zum Scaling-up-Ansatz* der GIZ aufweist (vgl. GIZ 2015a, S. 85f.): (1) die *projekt-/programmorientierte Nachhaltigkeit* ist dann gegeben, wenn andere Akteur*innen eingeführte Innovationen dauerhaft eigenständig weiterführen; (2) die *output- oder leistungsorientierte Nachhaltigkeit* ist dann gegeben, wenn andere Akteur*innen die eingeführte Innovation in ihrem eigenen Interesse und zum eigenen Nutzen dau-

erhaft übernehmen; (3) die *systemorientierte Nachhaltigkeit* beschreibt die *Breitenwirksamkeit*, d.h. die Veränderungen eines gesamten Systems, in dem eine Innovation beispielhaft eingeführt wurde, z.B. Reformen in einem Gesundheits- oder Bildungssystem; (4) bei der *innovationsorientierten Nachhaltigkeit* werden Innovationen übernommen und gemäß der Bedürfnisse von Akteur*innen modifiziert. Wird eine Maßnahme anhand der beschriebenen vier Dimensionen bewertet, entstehen sog. *Nachhaltigkeitsprofile*, die in einem *additiven Nachhaltigkeitsindex* zusammengefasst werden. Durch den Index kann eine *zusammenfassende Aussage zur Nachhaltigkeit* von Maßnahmen anhand von fünf Ausprägungen getätigt werden: Von *keiner Nachhaltigkeit*, sofern keine Dimension positiv bewertet wurde, bis hin zu einer *sehr hohen Nachhaltigkeit*, wenn die Maßnahme hinsichtlich aller Nachhaltigkeitsdimensionen positiv bewertet wurde (Caspari, 2004). Der Evaluationsleitfaden zum Ansatz (s. Tab. 8) ist v.a. auf die „Erfassung und Bewertung von Wirkungen ausgerichtet“, die durch Maßnahmen ausgelöst werden, wobei „Strukturen, Prozesse und Bedingungen [erfasst werden], unter denen Wirkungen entstehen“ (Stockmann, 2006, S. 179). Anhand des Leitfadens wird ersichtlich, dass der Ansatz auch Wirkungen und Nachhaltigkeit Meso- und Makroebene einbezieht, was in der SfE-Forschung bislang vernachlässigt wird.

Tab. 8: Struktur des Evaluationsleitfadens zum CEval-Ansatz (Stockmann, 2006, S. 183)

(1) Programm & Umwelt	(2) Programmverlauf
<ul style="list-style-type: none"> • Programmbeschreibung: Daten, Konzeption, Innovation, Ressourcen • Umwelt-/Kontextbedingungen: Praxis-/Politikfeld, Zielgruppenrelevanz 	<ul style="list-style-type: none"> • Vorbereitung, Planung • Steuerung • Vorbereitung des Förderendes • Nachbetreuung
(3) Interne Wirkungsfelder (Organisation)	(4) Externe Wirkungsfelder (Adressaten, Politikfelder)
<ul style="list-style-type: none"> • Zielakzeptanz • Qualifikationsniveau Personal • Organisationsstruktur • Finanzielle Ressourcen • Technologie • Interne Wirkungen 	<ul style="list-style-type: none"> • Zielakzeptanz bei Zielgruppen • Zielgruppenerreichung • Nutzen für Zielgruppen • Zielgruppenübergreifende Wirkungen • Wirkungen in Politikfeldern • Externe Wirkungen
(5) Programmqualität	
<ul style="list-style-type: none"> • Planungs- und Durchführungsqualität • Interne wirkungsbezogene Qualität • Externe wirkungsbezogene Qualität • Nachhaltigkeit auf der Meso-Ebene (Programmebene) • Nachhaltigkeit auf der Makro-Ebene (Effizienz, gesellschaftspolitische Relevanz, ökologische Verträglichkeit) 	

Darüber hinaus verdeutlichen die formulierten Schlussfolgerungen und Ausblicke die *zentrale Rolle von Evaluationsstandards*. Perspektivisch bieten sie die Möglichkeit, eine Grundlage für weitere Meta-Evaluationen zu sein, die die Qualität von Evaluationen erfassen. Nicht nur zur Qualitätssicherung der eigenen Evaluationen, sondern auch als „Ratgeber für die Evaluationspraxis“ (bspw. zur Beteiligung von Stakeholdergruppen, Wahrung der eigenen Unabhängigkeit) sollten Evaluierende die Standards nutzen (Widmer, 2000, S. 94). Evaluationspraktizierende bemängeln im Zusammenhang mit der Qualitätssicherung häufig die Vielzahl an Standards und eine damit einhergehende Schwierigkeit, sich mit und in ihnen zurechtzufinden. Eine *funktionale Übersichtsmatrix* der Standards hat den Vorteil, dass sie die Standards evaluativen Tätigkeiten wie bspw. der Berichtserstattung zuordnet (ebd.). Für Praktizierende wird hierdurch ersichtlich, i.R. welcher Tätigkeit auf welche Standards geachtet werden sollte. Anders als bei allgemeinen Listen zu den Standards (vgl. DeGEval, 2016a; DEval, 2018) können in der jeweiligen Evaluationsphase so die relevanten Standards überprüft werden und auch in der Kommunikation mit den Auftraggebenden nützlich sein. Da die DEval-Standards (2018) besonderen Wert auf die *Unabhängigkeit (Standards F1, F6)*, eine *genaue Umsetzungsplanung und das Monitoring (Standard N8)* legen, was die Umsetzung von Empfehlungen und damit auch die langfristige Nutzung von Evaluationen fördert, werden sie für die exemplarische Checkliste verwendet. In Grau hinterlegte Felder bedeuten, dass der Standard für diese Phase weniger relevant ist.

Tab. 9: Checkliste für Evaluierende zur Überprüfung von Evaluationsstandards im Evaluationsprozess (eigene Darstellung in Anlehnung an DEval, 2018; Widmer, 2000, S. 95)

Standards ³⁷¹	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	E 1	E 2	E 3	F 1	F 2	F 3	F 4	F 5	F 6	F 7	G 1	G 2	G 3	G 4	G 5	G 6	G 7	G 8	G 9	G 10	G 11	G 12	V 1	V 2		
Bereich	1	2	3	4	5	6	7	8	1	2	3	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8	9	10	11	12	1	2		
Vertragliche Vereinbarung																																		
Budget der Evaluation																																		
Evaluationspersonal																																		
Evaluationsmanagement																																		
Entscheidung zur Evaluation																																		
Definition des Problems																																		
Evaluationsdesign																																		
Datensammlung																																		
Datenauswertung																																		
Berichterstattung																																		
Kommunikation, Feedback																																		

³⁷¹ *Nützlichkeit*: N1 Politisch relevante Evaluierungsgegenstände, N2 Adressaten, Beteiligte, N3 Partizipation, N4 Klärung Evaluierungszwecke, N5 Eingrenzung des Evaluierungsgegenstands, N6 Nutzbarkeit, Offenlegung der Ergebnisse, N7 Rechtzeitigkeit, N8 Umsetzungsplanung, -monitoring; *Evaluierbarkeit*: E1 Durchführbarkeit einer Evaluierung, E2 Angemessene Verfahren, Methoden, E3 Effizienz der Evaluierung; *Fairness, Unabhängigkeit, Integrität*: F1 Unabhängigkeit der Evaluierung, F2 Berechenbarkeit, Planungssicherheit, F3 Ethisches Vorgehen, F4 Datenschutz, Datensparsamkeit, F5 Transparenz der Evaluierung, F6 Unparteiische, unabhängige Durchführung u. Berichterstattung, F7 Offenlegung von Werten; *Genauigkeit, Wissenschaftlichkeit, Nachvollziehbarkeit*: G1 Glaubwürdigkeit, Kompetenz des Evaluierungsteams, G2 Angemessene Beschreibung des Evaluierungsgegenstands, Zwecke, Fragestellungen, G3 Kontextanalyse, Beachten der Rahmenbedingungen, G4 Theorie, Evaluationsdesign, Methodenauswahl, G5 Nutzung u. Generierung valider, reliabler Informationen, Angabe von Informationsquellen, G6 Qualitätssicherung, systematische Fehlerprüfung, G7 Qualität der Datenanalyse, G8 Begründete Schlussfolgerungen, G9 Vollständigkeit der Berichterstattung, G10 Angemessener Umgang mit Kommentaren der Beteiligten, G11 Aufbewahrung(sfristen) von Daten, G12 Replizierbarkeit; *Vergleichbarkeit*: V1 Berücksichtigung einheitlicher Kriterien, Prüffragen, Bewertungsmaßstäbe, V2 Schaffung der Voraussetzungen für Querschnittsauswertungen.

Die DEval-Standards legen als einzige einen *Schwerpunkt auf die Umsetzung* und damit auf die nachhaltige Nutzung von Evaluationen. Da nach wie vor generell Schwierigkeiten bestehen, Evaluationsergebnisse in die Praxis zurückzuspielen und nachzuvollziehen, warum Handlungsempfehlungen als Teil von Evaluationsprozessen im Nachgang umgesetzt werden oder nicht, bietet es sich an, dass andere Forschungsarbeiten diesen Aspekt zum Gegenstand wählen.

8. Fazit und Ausblick: Interdisziplinarität und Gestaltungsspielräume als wegweisende Voraussetzungen

Basierend auf der Ausgangslage von SfE in Deutschland allgemein und dem Zugang der Autorin als wissenschaftliche Begleiterin zu den SfE-Maßnahmen der GIZ im Speziellen sowie resultierend aus den Forschungsdefiziten und den damit verbundenen -desideraten, hat sich die vorliegende Arbeit zum Ziel gesetzt, einen Beitrag (1) zur Verortung des Forschungsfeldes SfE, (2) zur Positionierung von Forschenden im Themenfeld und (3) zur Dokumentation der deutschen sportbezogenen EZ-Maßnahmen zu leisten.

Die Dokumentation erfolgte im zweiten Kapitel der Arbeit: Die Beschreibung der sportbezogenen Maßnahmen in Deutschland verdeutlicht, dass ‚der Sport‘ in der dt. EZ zwar früh Einzug gehalten hat, dies international jedoch aufgrund bislang fehlender Dokumentationen kaum bekannt ist. Durch die Dokumentation sollen nicht nur die langjährigen Bemühungen aufgezeigt werden, sondern auch eine Expertise, die nur mit wenigen Ländern vergleichbar ist und die nicht, wie es in den 1990er Jahren der Fall war, wieder in Vergessenheit geraten soll. Die Mittelzuwendungen an das AA für Sportförderung sind seit den 1960ern durchgehend, wenn auch leicht schwankend vorhanden, wohingegen sie an das BMZ für Maßnahmen zur *Entwicklung durch Sport* ab den 1970ern stark differenzieren, zeitweise ab Mitte der 2000er sogar ausgesetzt werden – während das Thema SfE in anderen Ländern als EZ-Medium einen besonderen Aufschwung erfährt. Erst ab 2013 findet SfE durch die Etablierung des SV in der GIZ einen klaren Einzug in die dt. EZ. Die zugewiesenen Fördermittel steigen sprunghaft an; durch das SV und weiterer regionaler Vorhaben werden SfE- sowie Baumaßnahmen in zahlreichen Ländern umgesetzt. Immer mehr und diversere Akteur*innen werden Partner*innen i.R. der Aktivitäten, die in den Ländern zu unterschiedlichen, sich an den SDGs orientierenden Ansätzen und meist für junge und marginalisierte Zielgruppen initiiert werden. Die Betrachtung der deutschen SfE-Maßnahmen, insbesondere der sportlichen Aktivitäten, vor dem Hintergrund pädagogischer und sportwissenschaftlicher Ansätze und Erklärungsmodelle zeigt, dass v.a. inhaltliche und funktionslogische Überschneidungen zur Erlebnispädagogik existieren. Es wird deutlich, dass die Zusammenarbeit mit anderen deutschen SfE-

Akteur*innen, v.a. mit Sport(fach)verbänden und Akteur*innen aus dem erlebnispädagogischen Bereich, maßgeblichen Einfluss auf die inhaltliche Ausgestaltung der Maßnahmen hat. D.h. die entwickelten Methoden in den Partnerländern beinhalten häufig Interaktionsspiele aus der Erlebnispädagogik oder haben vornehmlich klassische Mannschaftssportarten zum Gegenstand, wobei sportartenspezifische Übungen und Spiele im Training modifiziert oder kompetenzorientiert gestaltet werden.

Aus der terminologischen Reflexion resultiert, dass es trotz der Weiterentwicklung und Ausdifferenzierung des Begriffsverständnisses rund um *Sport* und *Entwicklung* in Deutschland im Laufe der Zeit bei der Bandbreite der aktuellen SfE-Maßnahmen der GIZ unklar bleibt, wie viele entwicklungspezifische Inhalte Teil der Maßnahmen sind oder ob eher eine (Breiten-)Sportförderung nach dem *Sport-für-Alle*-Prinzip und durch den Ausbau von Infrastruktur im Vordergrund stehen. Auch wenn eine erste begriffsbestimmende Einordnung von SfE bzw. von Sport für den *deutschen* EZ-Kontext in der Arbeit erfolgt ist, sollten weitere analytische Betrachtungen von SfE-Maßnahmen i.R. weiterer Forschungsarbeiten nicht nur im nationalen Kontext, sondern auch bzgl. anderer Länder und vergleichend erfolgen. Derartige Untersuchungen würden zu einer *definitiven* und damit auch zu einer *inhaltlichen Präzisierung* des Forschungsfeldes beitragen. Übergeordnet könnten nationale SfE-Ansätze und -Methoden miteinander verglichen und Unterschiede wie Gemeinsamkeiten, auch vor dem Hintergrund ‚nationaler Kulturexporte‘ herausgestellt werden. Forschungsleitende Fragen könnten bspw. folgende sein: Auf welche SDGs wird i.R. der Maßnahmen Bezug genommen? Beziehen sich diese auf die Zielgruppe (*People*) oder andere Prinzipien der Agenda 2030 (*Planet, Prosperity, Peace, Partnership*)? Welche Teilmaßnahmen finden über SfE-Projekte in Partnerländern statt (bspw. Baumaßnahmen, Lehrplanentwicklung, CD-Maßnahmen)? Welche Akteur*innen sind bei der Entwicklung der SfE-Methoden beteiligt und/ oder beeinflussen diese? Welche Sportart und sportpädagogischen Ausprägungen (bspw. sportdidaktisch, erlebnispädagogisch) beinhalten SfE-Methoden? Werden hierbei länder- und kulturspezifische Ansätze bedacht?

Das komplexe Untersuchungsmodell stößt *in seiner Ganzheit an Grenzen*, etwa bei der Betrachtung des SfE-Kontexts aus einer hauptsächlich aus der Evaluationsforschung geleiteten Perspektive. Teilaspekte der Rahmung, wie zur Qualität von Evaluationen, zur Evaluation von Wirkungen und Nachhaltigkeit sowie zur Rolle von Evaluierenden, decken jedoch bereits diagnostizierte Forschungsdefizite ab und könnten jeder für sich Gegenstand weiterer Forschungsarbeiten sein. Bedingt durch die Genese der Arbeit wurde ein monographischer Ansatz gewählt. Im Nachhinein wäre aus Sicht der Autorin eine *kumulative Vorgehensweise* von Vorteil gewesen, da Teile der Arbeit in Form englischsprachiger Artikel eine größere Forschergemeinde erreichen hätte. Ebenfalls der Entstehungsgeschichte der Arbeit geschuldet, kann im Hinblick auf die methodische Vorgehensweise nicht ausgeschlossen werden, dass erstens, die untersuchten Kontexte *wirklich in ihrer Gesamtheit* dargestellt werden und zweitens, die eigene Rolle als *involvierte Evaluierende und SfE-Forscherin* einen Einfluss auf die Darstellungen hat. Durch die theoretisch-konzeptionelle Rahmung können dennoch *Gegebenheiten, Besonderheiten und Defizite* im Hinblick darauf aufgezeigt werden, wie Evaluation bzw. Forschung im jeweiligen Kontext betrieben wird und welche Akteur*innen eine (beeinflussende) Rolle spielen:

Im Vergleich zu anderen Politikfeldern wurde in der Entwicklungspolitik bereits früh und besonders umfangreich evaluiert – mit dem Zweck, EZ-Maßnahmen zu steuern und zu kontrollieren. Erst später bekamen andere Evaluationsfunktionen mehr Gewichtung, wie etwa die Lern- oder Legitimationsfunktion und mit der Agenda 2030 v.a. die Aufklärungsfunktion. Die Vernetzung zwischen Evaluationseinheiten der EZ-Agenturen, die Etablierung von Arbeitsgruppen (bspw. die DAC Expert*innengruppe der OECD), nationalen und internationalen Evaluationsgesellschaften, Netzwerken und unabhängigen Institutionen führte zu einem intensiveren Austausch und zu einer zunehmenden Professionalisierung des Arbeitsbereichs. V.a. die OECD mit ihrer 1982 gegründeten Evaluations-Expert*innengruppe war federführend bei z.T. langwierigen Standardisierungsprozessen, sowohl was Evaluationstermini anbelangte, als auch im Hinblick von Ansätzen und Methoden. Bewährte theoriegeleitete (Logische Modelle) und partizipative Ansätze (PAR) wurden im Laufe der Zeit weiterentwickelt und modifiziert, wie bspw. in Form des WoM der GIZ. Auch multilaterale Organisationen (UN, Weltbank) waren u.a.

aufgrund der Tatsache, dass ihnen mehr finanzielle Ressourcen zur Verfügung standen, Vorreiter*innen, die die Gründung von internationalen Evaluationsnetzwerken ab den 2000ern vorantrieben. Mit der Einführung multidimensionaler Begriffe (z.B. Armutsmin-derung, Nachhaltigkeit) und immer komplexerer Zielsysteme (MDGs, SDGs) sowie der Verlagerung von einzelnen Projekten auf sektorenübergreifende Programme wurden auch die Anforderungen an die Evaluation und damit an Designs und Methoden immer komplexer. Den umfangreichen Zielsystemen und der Zunahme an zur Verfügung stehen- den Daten versuchen Forschende, mit quantitativen, aggregierbaren Standardindi- katoren zu begegnen. Auch wenn dadurch eine methodisch einseitige Fokussierung droht, ist die Nutzung eines quantitativ-qualitativen Methodenmix dennoch gängige Evaluationspraxis in der EZ. Es werden zunehmend umfangreiche Reviews im Allgemei- nen durchgeführt, um Ergebnisse mehrerer Studien zusammenfassend darzustellen o- der im Speziellen Studienergebnisse meta-analytisch auszuwerten, bspw. zu einzelnen SDGs. Nach wie vor spielen Meta-Evaluationen, die die Qualität von Evaluationen prü- fen, noch eine untergeordnete Rolle. Anspruchsvolle Designs wie RCTs sind immer häu- figer gefragt, v.a., um Wirkungen von Maßnahmen nachzuweisen. Die Debatte um die Wirksamkeit von EZ wird seit den 1970er Jahren durchgehend geführt, durch internati- onale Konsense (bspw. Paris Deklaration von 2005) verstärkt – sie bestimmt(e) die EZ- Evaluation maßgeblich. Die Nachhaltigkeit von Maßnahmen wird i.R. von Evaluations- kriterien hauptsächlich im Sinne von Dauerhaftigkeit bewertet und greift zu kurz. Ex- post Evaluationen, die die Nachhaltigkeit und Wirkungen von Maßnahmen untersuchen, werden zu wenig veranlasst. Interdisziplinär vorangetriebene Fortschritte und Entwick- lungen in der Evaluationsforschung, die sich der mit der Agenda 2030 verbundenen Her- ausforderungen annehmen, sind in der Praxis noch nicht angekommen. Aufgrund der Kritik an der mangelnden Unabhängigkeit von Evaluation wurden in zahlreichen EZ- Agenturen interne unabhängige Abteilungen oder Einheiten etabliert, externe Evalua- tionsinstitutionen gegründet und vermehrt externe Gutachtende und Beratungsfirmen mit Evaluationen beauftragt. Es hat sich ein Markt für Evaluation entwickelt, der letzt- endlich in Bezug auf die Evaluierenden Konkurrenz und Abhängigkeiten generiert.

Im Vergleich zu anderen Forschungsfeldern ist jenes zu SfE noch ‚jung‘. Seit seiner Etablierung im internationalen Kontext ab der Jahrtausendwende hat es sich enorm weiterentwickelt: Sowohl in der Quantität als auch in der Qualität von Studien wurden Fortschritte erzielt, was nicht zuletzt darauf zurückzuführen ist, dass es von zahlreichen Forschenden aus unterschiedlichen Disziplinen bearbeitet wird. Dennoch bestehen nach wie vor große Herausforderungen: Die unzureichende Abgrenzung der Termini *research* und *evaluation*, die zweidimensionale Funktionszuschreibung *accountability* und *learning* in Anlehnung an geldgebende und -empfangende Akteur*innen sowie die Spaltung in Verfechter*innen theoriegeleiteter-quantitativer bzw. partizipativer-qualitativer Ansätze und Methoden führt zu *polarisierenden und teilweise undifferenzierten Debatten* unter den Forschenden. Die Vielzahl der beteiligten Wissenschaftler*innen aus diversen Disziplinen bedingt die Nutzung unterschiedlichster theoretischer Modelle, identifizierte SfE-Schlüsselkonzepte werden wenig genutzt. Auch im Kontext von SfE nehmen Reviews zu, Meta-Evaluationen oder als ex-post Evaluationen ausgewiesene Studien können nicht identifiziert werden. Die unzureichende Nutzung bewährter Ansätze und Modelle aus der Entwicklungsforschung und der EZ-Evaluation wird zwar angemahnt, bislang jedoch von wenigen Forschenden umgesetzt, wie bspw. Lindsey i.R. einer Nachhaltigkeitskonzeption. Fehlende Studien auf Meso- und Makro-Ebene sowie fehlende Langzeit- und Vergleichsstudien, auch mit dem Fokus auf Wirkungen und Nachhaltigkeit, werden v.a. auf mangelnde finanzielle Ressourcen, die Problematik der Etablierung im Wissenschaftssystem und auf die *Fragmentierung bzw. Multidisziplinarität* des Forschungsfeldes zurückgeführt. Dadurch, dass nationale und internationale EZ-Akteur*innen den Sport als kostengünstiges und effizientes Medium zur Erreichung politischer Zielsetzungen erkannt haben und zahlreiche Maßnahmen und damit auch lokale Partnerorganisationen finanzieren, haben sie großen Einfluss auf das Forschungsfeld: Erstens spielen EZ-Akteur*innen (UN, Commonwealth Secretariat, GIZ) eine große Rolle bei der inhaltlichen Standardisierung von Ansätzen und Methoden; zweitens ist der Druck, Wirkungen des „low-cost, high-impact tools“ über Evaluationen nachzuweisen, besonders groß; und drittens ermöglichen ihre finanziellen Zuwendungen maßgeblich die Durchführung von Studien durch externe Wissenschaftler*innen aus dem Globalen Norden. Diese Abhängigkeit bedingt, dass das Forschungsfeld und auch die Forschenden

sich in einem *Spannungsfeld zwischen Verwissenschaftlichung und Anwendungsorientierung* befinden.

Die wissenschaftliche Begleitung des SV SfE der GIZ als praktisches Beispiel zur Evaluation im SfE-Kontext zeigt, dass nach anfänglicher Orientierung die Evaluationsprozesse nach bewährten Ansätzen und Methoden der EZ gestaltet wurden (WoM, Capacity Works, Mixed Methods). Die meisten Evaluationen nutzten eher quantitativ-angelegte Fragebögen im Vorher-Nachher-Vergleich und wurden prozessbegleitend (on-going) auf Projektebene (Mikro- und Meso-Ebene) durchgeführt. Nur eine Evaluation fand im Nachgang von beendeten Projekten (ex-post) statt. Neben der DSHS als externe Evaluierende war die GIZ (intern) an fast allen Studien beteiligt. Weitere Akteur*innen wie Partnerorganisationen waren/sind hauptsächlich i.R. der Wirkungsevaluationen eingebunden, häufig bei der Entwicklung von Instrumenten und der Sammlung von Daten. Die Frage nach Wirkungen auf der Ebene von Kindern und Jugendlichen wurde früh gestellt. Bereits in den ersten beiden Phasen wurden diesbezüglich Evaluationen initiiert und Erfahrungen sowie organisatorisch-methodische Herausforderungen in einem Grundsatzpapier festgehalten. Evaluationsstandards wurden zwar nicht als Referenz verwendet, die meisten Genauigkeitsstandards konnten jedoch eingehalten werden. Weiterführend wäre hier von Interesse, wie praktische Beispiele in anderen Ländern aussehen, also was Bestandteile und Schwerpunkte der wissenschaftlichen Begleitung von anderen EZ-Akteur*innen sind. Voraussetzung hierfür wären *Vertragsbedingungen*, die eine detaillierte Beschreibung der (Evaluations-)Aktivitäten inklusive Ergebnisdarstellungen und eine Auswertung im Sinne einer Meta-Evaluation zuließen sowie eine umfangreiche *Reflexion der Evaluierenden*, was im Falle der wissenschaftlichen Begleitung des SV SfE der GIZ so nicht möglich war.

Durch den evaluierenden Blick auf die Theorie und Praxis der drei Kontexte können Kernaussagen abgeleitet werden, die *Defizite, Herausforderungen und Möglichkeiten* erfassen und letztendlich zu den ersten zwei Zielsetzungen der Arbeit beitragen: (1) Der Verortung des Forschungsfeldes SfE sowie (2) der Positionierung von Forschenden im SfE-Kontext.

Damit sich das Forschungsfeld aus dem Spannungsverhältnis lösen kann, sind Veränderungen sowohl auf Seiten der Forschenden als auch auf Seite der EZ-Akteur*innen notwendig. Für Forschende bedeutet dies: Wenn sie in ihrer disziplinären Isolation bleiben und keine gemeinsame ‚neue Forschungsagenda‘ verfolgen, für die bereits erste Grundsteine gelegt sind (vgl. Giulianotti et al., 2019), wird das Feld in einem verwirrenden Theoretisierungsprozess ohne Aussicht auf Weiterentwicklung stecken bleiben – und letztlich auch ohne Mehrwert für SfE-Organisationen sein, die auf Forschung angewiesen sind, um ihre Arbeit zu verbessern und gegenüber Geldgebenden zu legitimieren. Eine Weiterentwicklung ist möglich, wenn es gelingt, sich *von einem multidisziplinären zu einem interdisziplinären Forschungsfeld* zu entwickeln. Um dies zu erreichen, müssen eine gemeinsame Sprache (v.a. bzgl. Definitionen) und gemeinsame Problemsichten (Orientierung an den SDGs) gefunden, und durch den Blick in andere Forschungs- und Evaluationsbereiche (v.a. in die allgemeine EZ) theoretische Ansätze und Methoden vereinheitlicht werden.

Auf Seiten der EZ-Akteur*innen heißt dies, dass Evaluationen wünschenswerterweise nicht mehr nur vornehmlich im Sinne der Legitimation und Kontrolle durchgeführt werden, auch wenn diese nach wie vor eine große Rolle spielen, vor allem wenn es um die Legitimation von Maßnahmen durch qualitativ hochwertige Wirkungsevaluationen geht. Hoffnung liegt in der *Evaluationsfunktion der gesellschaftlichen Aufklärung*, die zunehmend an Bedeutung gewinnt und die nur stattfinden kann, wenn sich Evaluation *aus der Zwangsjacke der Wünsche durch die Auftraggebenden* befreit, d.h. seitens der Evaluierenden freier und flexibler gestaltet werden kann.

Schlussendlich kann das Forschungsfeld nur dann Bestand haben, wenn sich *stabilere Finanzierungsstrukturen* entwickeln, die bislang weder über das EZ- noch über das Hochschulsystem gegeben sind. Optimistisch kann in beiden Fällen mit der *Nachhaltigkeitsagenda* argumentiert werden, die auf beide Systeme auch künftig einen großen Einfluss haben wird und i.R. derer Zielerreichung der Sport eine umfangreiche Rolle spielen kann. Über den Beitrag zu den SDGs hinaus werden über SfE-Maßnahmen immer ‚neue‘ weltpolitische Themen angegangen: Fluchtursachenbekämpfung, Klima und auch die

Coronapandemie. Mit einem Fokus auf COVID-19 hat die UNESCO erst Ende vergangenen Jahres ihr „sport-based flagship“ mit dem Titel *Fit for Life* aufgesetzt, was zeigt, dass große bilaterale EZ-Organisationen das Thema ‚nicht aufgeben‘. In zwei Phasen soll bis 2032 der Sport zur Förderung von Gesundheit, Bildung und Gleichstellung eingesetzt werden. Die Programmaktivitäten verbinden internationale, regionale und nationale Akteure, darunter auch die GIZ, „um die Erholung von COVID-19 zu beschleunigen und das Wohlbefinden aller Menschen, angefangen bei der Jugend, zu verbessern“. „Building evidence“ über M&E-Maßnahmen spielen hierbei ebenfalls eine zentrale Rolle (UNESCO, 2021).

Mit pessimistischem Blick kann angenommen werden, dass die Erhaltung des Arbeits- und Forschungsfeldes letztendlich an der *Motivation und Überzeugung von Einzelpersonen* hängt, seien es Forschende, die vom Mehrwert von SfE überzeugt sind, oder sportbegeisterte Entwicklungsminister*innen. Im Falle der neuen Entwicklungsministerin Svenja Schulze, die in ihrer Jugend „völlig sportbegeistert“ war, besteht Hoffnung. Zudem gibt sie an, dass der Sport sie „politisiert“ habe, da sie sich in diesem Rahmen erstmalig für Gleichberechtigung eingesetzt hatte (Schulze, 2021). Ob das auch heißt, dass der Sport wieder Erwähnung im Entwicklungsbericht der Bundesregierung finden wird, bleibt abzuwarten.

Es wird prognostiziert, dass *Spannungen zwischen Wissenschaft und Praxis*, demnach auch zwischen einer eher verwissenschaftlichen oder anwendungsorientierten Vorgehensweise i.R. von Evaluationsaktivitäten, *zunehmen* werden. Die Entwicklungen hängen laut Evaluationsforschenden wesentlich von der Kultur der Auftraggebenden und der Marktentwicklung ab (vgl. Meyer & Stockmann, 2017). Die *Marktentwicklung* kann seitens der SfE-Forschenden letztendlich nicht beeinflusst werden, auch wenn angenommen wird, dass Sport im EZ-Kontext nicht verschwindet, die Frage nach der Wirkung von Sport eine zentrale bleibt und damit auch der Bedarf an Evaluationen weiterhin vorhanden ist. Für SfE-Forschende gilt allerdings, sich eingehender mit den Evaluierungspolicies, also mit den *Evaluationskulturen der Auftraggebenden* auseinandersetzen, die sehr unterschiedlich sein können: Werden Ressourcen für qualitativ hochwertige Studien investiert oder reicht eine ‚Story from the field‘ aus, um die Auftraggebenden zu

befriedigen? Darüber hinaus sollten sich SfE-Forschende an *Evaluationsstandards* orientieren, die eine Orientierung hinsichtlich der *Einbindung von Beteiligten (Partizipation)* sowie der *Objektivität* und *Unabhängigkeit* der Forschenden bietet, bspw. in Form von Checklisten. Zudem sollten sie i.R. ihrer Forschungsaufgaben stets eine *reflexive Haltung* im Hinblick auf ihre eigene Identität und ihre Rollen einnehmen, um sich auch in der besonderen Beziehung zwischen Politik und Evaluation positionieren zu können. Für SfE-Forschende bleibt die Frage, wie und wo sich letztendlich verorten werden und wollen. Verstehen sie sich mehr als *Dienstleistende mit Servicefunktion* oder als *qualitätsorientierte Wissenschaftler*innen*? Welche Kompromisse gehen sie ein? Die Entscheidung ist eine essenzielle, aber individuelle.

Literaturverzeichnis

- 3IE. (2019). *3iE at a Glance*. Zugriff am 16. Oktober 2019 unter <https://www.3ieimpact.org/about-us/performance-metrics>
- 3IE. (2021). *Evidence gap maps*. Zugriff am 12. März 2021 unter <https://www.3ieimpact.org/evidence-hub/evidence-gap-maps>
- AA. (2018, 07. März). *Sport ohne Grenzen: Die internationale Sportförderung des Auswärtigen Amtes*, Auswärtiges Amt. Zugriff am 23. Oktober 2020 unter <https://kampala.diplo.de/ug-de/themen/kultur/-/1679274>
- AA. (2018, 09. Mai). *Sport baut Brücken: Deutsche Fußballbotschafter ausgezeichnet*, Auswärtiges Amt. Zugriff am 23. Oktober 2020 unter <https://www.auswaertiges-amt.de/de/aussenpolitik/themen/kulturdiallog/-/2075232>
- Abbasi, S. (2019, 25. November). *International framework on measuring sport's contribution gains traction*. Zugriff am 20. Mai 2020 unter <https://thecommonwealth.org/media/news/international-framework-measuring-sport%E2%80%99s-contribution-gains-traction>
- Abdel-Malek, T. (2015). *The Global Partnership for Effective Development Cooperation. Origins, actions and future prospects* (Studies / Deutsches Institut für Entwicklungspolitik, 88). Bonn: Deutsches Institut für Entwicklungspolitik (DIE).
- Abma, T. A. (2007). The Social relations of Evaluation. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of evaluation. Policies, programs and practices* (pp. 184-199). London: Sage Publ.
- Achleitner, A.-K. (2018, 14. Februar). *Social Entrepreneurship. Definition: Was ist "Social Entrepreneurship"?* Gabler Wirtschaftslexikon. Zugriff am 13. Oktober 2021 unter <https://wirtschaftslexikon.gabler.de/definition/social-entrepreneurship-52240/version-275382>
- Adams, A., Harris, K. & Lindsey, I. (2018). Examining the capacity of a sport for development programme to create social capital. *Sport in Society*, 21 (3), 558-573. doi: 10.1080/17430437.2017.1346627.
- AEA. (2018). *Guiding Principles for Evaluators*, American Evaluation Association. Zugriff am 21. März 2020 unter <https://www.eval.org/p/cm/ld/fid=51>
- AfrEA. (2006/2007). *African Evaluation Guidelines - Standards and Norms* (African Evaluation Association (AfrEA), Hrsg.). Zugriff am 24. November 2019 unter https://member.samea.org.za/uploads/docs/AEG_2007_EN.pdf
- Aiazzi, T. (2016a). *Evaluation in the SDG era: lessons, challenges and opportunities for UNEG. Volume I* (UNEG Publications and Report). Zugriff am 21. Januar 2020 unter <http://www.uneval.org/document/download/2588>
- Aiazzi, T. (2016b). *Evaluation in the SDG era: lessons, challenges and opportunities for UNEG. Volume II* (UNEG Publications and Report). Zugriff am 21. Januar 2020 unter <http://www.unevaluation.org/document/download/2589>
- Alkin, M. & Christie, C. (2009). An Evaluation Theory Tree. In M. C. Alkin (Hrsg.), *Evaluation roots. Tracing theorists' views and influences* (S. 13-65). Thousand Oaks: Sage Publ.

- Ambrosy, R. (2014). *Die Finanzierung der deutschen Hochschulen*. Oldenburg: Carl von Ossietzky Universität Oldenburg. Zugriff am 19. Dezember 2021 unter https://www.nrw.uni-kanzler.de/fileadmin/user_upload_NRW/Publikationen/2014_Ambrosy_Die_Finanzierung_der_deutschen_Hochschulen.pdf
- Andresen, R., Rieder, H. & Trosien, G. (1989). Sportentwicklungspraxis, -politik und -forschung in der Bundesrepublik Deutschland. oder: Von der nationalen Selbstdarstellung zur Entwicklungszusammenarbeit im Sport? In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 9-17). Schorn-dorf: Hofmann.
- Ashoff, G. (2010). Wirksamkeit als Legitimationsproblem und komplexe Herausforderungen der Entwicklungspolitik. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 27-68). Baden-Baden: Nomos.
- Ashoff, G. (2011). *Was hat die Paris-Erklärung zur Wirksamkeit der Entwicklungszusammenarbeit gebracht? Ergebnisse der internationalen Evaluierung und Schlussfolgerungen für die deutsche EZ* (Analysen und Stellungnahmen 16/2011). Deutsches Institut für Entwicklungspolitik (DIE).
- Ashoff, G. (2013). *50 Jahre entwicklungspolitische Länderprüfungen (Peer Reviews) der OECD: ein Instrument der Qualitätssicherung und des gemeinsamen Lernens* (Analysen und Stellungnahmen 4/2013). Deutsches Institut für Entwicklungspolitik (DIE).
- Astbury, B. (2017). Von der Evaluationstheorie zum Testen von Evaluationstheorien? In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 223-242). Münster: Waxmann Verlag GmbH.
- Baig-Schneider, R. (2012). *Die moderne Erlebnispädagogik. Geschichte, Merkmale und Methodik eines pädagogischen Gegenkonzepts* (Praktische Erlebnispädagogik). Augsburg: ZIEL.
- Baker, R. E., Baker, P. H., Atwater, C. & Andrews, H. (2015). Sport for development and peace: a program evaluation of a sport diplomacy initiative. *International Journal of Sport Management and Marketing*, 16 (1/2), 52. doi: 10.1504/IJSMM.2015.074932.
- Balz, E. & Kuhlmann, D. (2009). *Sportpädagogik. Ein Lehrbuch in 14 Lektionen* (3. Auflage). Aachen: Meyer & Meyer.
- Bamberger, M., Segone, M. & Tateossian, F. (2016). *Evaluating the Sustainable Development Goals. With a "No one left behind" lens through equity-focused and gender-responsive evaluations*. Zugriff am 20. Januar 2020 unter <https://evalpartners.org/sites/default/files/documents/evalgender/EN-Evaluating-SDG-web.pdf>
- Bamberger, M., Vaessen, J. & Raimondo, E. (Eds.). (2015). *Dealing with complexity in development evaluation. A practical approach*. Los Angeles: Sage.

- Barisch-Fritz, B. & Volk, C. (2016). Übersicht zum Thema Interdisziplinarität und Sportwissenschaft. *Ze-phir*, 23 (2), 4-7.
- Bauer, M. (2004). *Entwicklungszusammenarbeit und Breitensport. Theoretische Überlegungen zum Breitensport in der Entwicklungszusammenarbeit und Evaluierung von Breitensportveranstaltungen in Entwicklungs-, Schwellen- und Transformationsländern innerhalb der Förderstrategie des Internationalen Olympischen Komitees*. Zugl.: Würzburg, Univ., Diss., 2002. Frankfurt am Main: Lang.
- Baur, N., Kelle, U. & Kuckartz, U. (Hrsg.). (2017). *Mixed methods* (Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderheft, Bd. 57). Wiesbaden: Springer VS.
- Berlage, L. & Stokke, O. (1992). *Evaluating Development Assistance - Approaches and Methods*. London: Frank Cass.
- Bethge, J. P., Steurer, N. & Tscherner, M. (2011). Nachhaltigkeit. Begriff und Bedeutung in der Entwicklungszusammenarbeit. In J. König & J. Thema (Hrsg.), *Nachhaltigkeit in der Entwicklungszusammenarbeit. Theoretische Konzepte, strukturelle Herausforderungen und praktische Umsetzung* (Globale Gesellschaft und internationale Beziehungen, 1. Aufl., S. 15-40). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Beyer, E. & Friedel, P. (1979). *Sportförderung in Ländern der Dritten Welt* (GTZ Schriftenreihe, 85). Eschborn: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ).
- Beywl, W. (2020). Rezension zu: Patton, Michael Q.: *Blue Marble Evaluation. Premises and Principles*. *Zeitschrift für Evaluation (ZfEv)*, 19 (2), 343-348.
- Beywl, W. & Widmer, T. (2009). Evaluation in Expansion: Ausgangslage für den intersektoralen Dreiländer-Vergleich. In T. Widmer, W. Beywl & C. Fabian (Hrsg.), *Evaluation. Ein systematisches Handbuch* (1. Aufl., S. 13-23). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Biermann, M. (2016). *“...the Power to Change the World”? Analysis of Sport-in-Development Programmes in Khayelitsha, South Africa and its Challenges for Research(ers)*. Dissertation, Deutsche Sporthochschule Köln.
- Binnendijk, A. (2000). *Results Based Management in the Development Co-operation Agencies: A Review of Experiences. Background Report*. Paris: OECD, DAC Working Party on Aid Evaluation. Zugriff am 20. Juni 2019 unter <https://www.oecd.org/dac/evaluation/dcdndep/31950852.pdf>
- Birdsall, K. (2019). *Growing Together. How volunteerism fosters social cohesion between Syrian and Turkish young people*. Bonn, Eschborn: GIZ GmbH.
- Birnbrich, T. (Oktober 2021). *Zusammen nachhaltig: ‘Sport for Development’ wins 1. FC Köln for Cooperation in Sustainability*. Spo(r)tlight on! Newsletter: 6. Zugriff am 19. Januar 2022 unter <https://www.sport-for-development.com/im-glib/newsletter/2021-giz-en-sporlight-on-newsletter-no6-q3.html>
- Black, D. R. (2017). The challenges of articulating ‘top down’ and ‘bottom up’ development through sport. *Third World Thematics: A TWQ Journal*, 2 (1), 7-22. doi: 10.1080/23802014.2017.1314771.

- Black, D. R. (2019). SDP and global development. In H. Collison, S. C. Darnell, R. Giulianiotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 35-45). Milton: Routledge.
- Blessing-Kapelke, U., Klages, A. & Schwind-Gick, G. (2007). *Demographische Entwicklung in Deutschland: Herausforderung für die Sportentwicklung. Materialien, Analysen, Positionen* (Projektbericht der Projektgruppe „Demographischer Wandel“ des Deutschen Olympischen Sportbundes (2005 bis 2007), Hrsg.). Frankfurt a.M.: DOSB.
- Bliss, F. (2006). Evaluierungen in der Entwicklungszusammenarbeit. Ziele, Kriterien und Standards. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 13-40.
- Bliss, F., Neubert, S. & Heinz, M. (2007). Partizipativ gestaltete Evaluationen: Ein Diskussionspapier. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 161-228.
- BMZ. (2000). *Langfristige Wirkungen deutscher Entwicklungszusammenarbeit und ihre Erfolgsbedingungen* (BMZ-Spezial Nr. 19), Bonn.
- BMZ. (2006). *Evaluierungskriterien für die deutsche bilaterale Entwicklungszusammenarbeit. Eine Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen* (BMZ, Hrsg.). Zugriff am 24. Oktober 2018 unter https://www.bmz.de/de/zentrales_downloadarchiv/erfolg_und_kontrolle/evaluierungskriterien.pdf
- BMZ. (2008). *BMZ Konzepte 165 - Leitlinien für die bilaterale Finanzielle und Technische Zusammenarbeit mit Kooperationspartnern der deutschen Entwicklungszusammenarbeit* (BMZ, Hrsg.), Bonn, Berlin. Zugriff am 25. Januar 2019 unter <https://www.bmz.de/de/mediathek/publikationen/archiv/reihen/strategiepa-piere/konzept165.pdf>
- BMZ. (2010-2020a). *Deutsches Engagement: Sport als übergreifendes Thema für Entwicklung nutzen*. Zugriff am 03. Februar 2020 unter http://www.bmz.de/de/themen/sport-fuer-entwicklung/deutsches_engagement/index.html
- BMZ. (2010-2020b). *Fluchtursachen mindern – Aufnahmeregionen stabilisieren – Flüchtlinge unterstützen. BMZ-Sonderseite zum Thema Flucht*. Zugriff am 22. Oktober 2020 unter <https://www.bmz.de/webapps/flucht/index.html#/de>
- BMZ. (2010-2020c). *Evaluierungsberichte, Berichte zu Reviews und Wirkungsmessungen*. Zugriff am 13. März 2020 unter <https://www.bmz.de/de/ministerium/evaluierung/ergebnisse/index.html>
- BMZ. (2010-2019). *Index der menschlichen Entwicklung (HDI)*. Zugriff am 25. November 2019 unter https://www.bmz.de/de/service/glossar/l/index_hdi.html
- BMZ. (Mai 2013). *14. Entwicklungspolitischer Bericht der Bundesregierung. Weißbuch zur Entwicklungspolitik* (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Hrsg.), Bonn, Berlin.
- BMZ. (2015). *Die Rolle des Sports in der deutschen Entwicklungszusammenarbeit* (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ),

- Hrsg.). Zugriff am 25. Oktober 2020 unter http://www.bmz.de/de/mediathek/publikationen/reihen/infobroschueren_flyer/infobroschueren/Materiale268_sport_ez.pdf
- BMZ. (März 2017). *Entwicklungspolitik als Zukunfts- und Friedenspolitik. 15. Entwicklungspolitischer Bericht der Bundesregierung* (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Hrsg.), Bonn, Berlin.
- BMZ. (Oktober 2017). *Handbuch der bilateralen EZ. Verfahrensablauf für die Planung und Durchführung von Maßnahmen der FZ und TZ* (Handreichung Nr. HR023), Internes Dokument.
- BMZ. (2018). *Entwicklungspolitik 2030. Neue Herausforderungen - neue Antworten*. BMZ Strategiepapier, Bonn. Zugriff am 13. Januar 2020 unter https://www.bmz.de/de/mediathek/publikationen/reihen/strategiepa-piere/Strategiepapier455_06_2018.pdf
- BMZ. (2018, 26. März). *Gerald Asamoah und Nia Künzer sind neue Botschafter "Fußball für Entwicklung"*. Zugriff am 25. Oktober 2020 unter https://www.bmz.de/de/presse/aktuelleMeldungen/2018/mar-erz/180326_pm_019_Gerald-Asamoah-und-Nia-Kuenzer-sind-neue-Botschaf-ter-Fussball-fuer-Entwicklung/index.jsp
- BMZ. (2018, 11. Mai). *Projektlaunch: "Athletics for Development" in Uganda*. Zugriff am 25. Oktober 2020 unter http://www.bmz.de/de/presse/aktuelleMeldun-gen/2018/mai/180511_Projektlaunch-Athletics-for-Development-in-Uganda/in-dex.jsp
- BMZ. (2021). *Grundsätze der Evaluierung, Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit. Orientierungslinien des BMZ*. Zugriff am 11. März 2022 unter <https://www.bmz.de/re-source/blob/92894/3e098f9f4a3c871b9e7123bbef1745fe/evaluierungskrite-rien.pdf>
- BMZ. (September 2021). *EINEWELT – unsere Verantwortung. Globalisierung gerecht gestalten. 16. Entwicklungspolitischer Bericht der Bundesregierung* (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Hrsg.), Bonn, Berlin.
- Bne-portal. (o.D.). *Bildung für Nachhaltige Entwicklung: Hochschule*. Zugriff am 19. Dezember 2021 unter <https://www.bne-portal.de/bne/de/nationaler-aktions-plan/die-bildungsbereiche-des-nationalen-aktionsplans/hochschule/hochschule.html>
- Böhnke, J. (2010). *Abenteuer- und Erlebnissport. Ein Handbuch für Schule, Verein und Jugendsozialarbeit* (2. Aufl.). Münster: LIT-Verlag.
- Böhret, C., Jann, W. & Kronenwett-Löhrlein, E. (1988). *Innenpolitik und politische Theorie. Ein Studienbuch*. Opladen: Westdt. Verlag.
- Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. *Systematic reviews*, 5, 74. doi: 10.1186/s13643-016-0249-x.

- Borrmann, A. (1999). Evaluating development co-operation. The German case. *Intereconomics*, 34 (5), 255-259. doi: 10.1007/BF02927277.
- Borrmann, A., Fasbender, K., Holthus, M., Gleich, A. v., Reichl, B. & Shams, R. (1999). *Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Analyse, Bewertung, Reformen; Studie im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung* (1. Aufl.). Baden-Baden: Nomos.
- Borrmann, A., Gleich, A. v., Holthus, M. & Shams, R. (2001). *Reform der Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Eine Zwischenbilanz; Studie im Auftrag des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung* (HWWA-Studien, 63, 1. Aufl.). Baden-Baden: Nomos.
- Borrmann, A. & Stockmann, R. (2009a). *Evaluation in der deutschen Entwicklungszusammenarbeit. Band 1 - Systemanalyse*. Münster: Waxmann.
- Borrmann, A. & Stockmann, R. (2009b). *Evaluation in der deutschen Entwicklungszusammenarbeit. Band 2 - Fallstudien*. Münster: Waxmann.
- Böttcher, W., Kerlen, C., Maats, P., Schwab, O. & Sheikh, S. (Hrsg.). (2014). *Evaluation in Deutschland und Österreich. Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval - Gesellschaft für Evaluation*. Münster: Waxmann.
- Brand, A. (2020). Sport und globale Zusammenarbeit: Von Hoffnungen, Enttäuschungen und Erfolgen. In K. Petry (Hrsg.), *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage, S. 103-118). Leverkusen: Verlag Barbara Budrich.
- Brand, R. (2006). Mehr Worte als Taten? Der deutsche Beitrag zur Erfüllung der Millennium-Entwicklungsziele. In F. Nuscheler & M. Roth (Hrsg.), *Die Millennium-Entwicklungsziele. Entwicklungspolitische Königsweg oder ein Irrweg?* (Eine Welt, N.F., 20, S. 61-80). Bonn: Dietz.
- Brandt, T. (2007). Sozialer Kontext der Evaluation. In R. Stockmann (Hrsg.), *Handbuch zur Evaluation. Eine praktische Handlungsanleitung* (Sozialwissenschaftliche Evaluationsforschung, 6, S. 164-194). Münster: Waxmann.
- Brown, A. & Foster, M. (Januar 2001). *The Status of sector-wide Approaches* (Working Paper 142). London: ODA, Centre for Aid and Public Expenditure
- Bruder, M., Faust, J. & Krämer, M. (2019). *Rigorose Wirkungsevaluierung in der deutschen Entwicklungszusammenarbeit*. DEval Policy Brief 5/2019: DEval
- Burnett, C. (2008). Participatory action research (PAR) in monitoring and evaluation of sport-for-development programmes. *African Journal for Physical, Health Education, Recreation and Dance*, 14 (3). doi: 10.4314/ajpherd.v14i3.24805.
- Burnett, C. (2009). Engaging sport-for-development for social impact in the South African context. *Sport in Society*, 12 (9), 1192-1205. doi: 10.1080/17430430903137852.
- Burnett, C. (2012). *Stories from the field. YDF footprint in Africa*. Johannesburg: 30° South.

- Burnett, C. (2013a). The 'uptake' of a sport-for-development programme in South Africa. *Sport, Education and Society*, 20 (7), 819-837. doi: 10.1080/13573322.2013.833505.
- Burnett, C. (2013b). GIZ/YDF and youth as drivers of Sport for Development in the African context. *Journal of Sport for Development*, 1 (1), 4-15.
- Burnett, C. (2014). The impact of a sport-for-education programme in the South African context of poverty. *Sport in Society*, 17 (6), 722-735. doi: 10.1080/17430437.2014.882903.
- Burnett, C. (2015). Assessing the sociology of sport: On Sport for Development and Peace. *International Review for the Sociology of Sport*, 50 (4-5), 385-390. doi: 10.1177/1012690214539695.
- Burnett, C. & Hollander, W. J. (2007). The Sport-in-Development Impact Assessment Tool (S•DIAT). *African Journal for Physical, Health Education, Recreation and Dance*, June, 123-135.
- Busan HLF. (2011, Dezember). *Busan Partnership for Development Co-operation. Fourth High Level Forum on Aid Effectiveness, Busan, Republic of Korea, 29. November - 1. December 2011*. Zugriff am 15. Januar 2020 unter http://effectivecooperation.org/wp-content/uploads/2016/03/OUTCOME_DOCUMENT_-_FINAL_EN.pdf
- Bussmann, W., Klöti, U. & Knoepfel, P. (Hrsg.). (1997). *Einführung in die Politikevaluation*. Basel: Helbing und Lichtenhahn.
- Cailloux, M. (1988). Evaluation in CIDA and other Donors: Past and Future. In A. Kuyvenhoven & M. van Pelt (Eds.), *Evaluating project performance methodology and development of capabilities. Proceedings of a seminar organised by the Directorate-General for International Co-operation of the Netherlands Ministry of Foreign Affairs and the Department of the Auditor-General of Pakistan, 10-11 June 1987, The Hague* (pp. 155-164). The Hague.
- CampGroup. (2015, 28. Oktober). *Mit Ballsport für Mädchenrechte in Afghanistan*. Zugriff am 18. Januar 2022 unter <http://camp-group.org/camp-group-beraet-die-giz-in-kabul-afghanistan/>
- CARE, Mercy Corps, Schwery Consulting. (2008). *Moving Forward Toolkit. A Guide for Practitioners in the Field of Sport for Youth Emergencies* (International Network of Sport and Development Consultants (INSDC), Hrsg.). Zugriff am 19. April 2020 unter https://www.sportanddev.org/sites/default/files/downloads/part_i_toolkit_manual.pdf
- Caspari, A. (2004). *Evaluation der Nachhaltigkeit von Entwicklungszusammenarbeit. Zur Notwendigkeit angemessener Konzepte und Methoden* (Sozialwissenschaftliche Evaluationsforschung, 3). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Caspari, A. (2008). (Rigorous) Impact Evaluations - Eine nicht nur für die Entwicklungszusammenarbeit relevante internationale Diskussion. *Zeitschrift für Evaluation (ZfEv)*, 7 (1), 137-141.

- Caspari, A. (2009). 'Rigorese' Wirkungsevaluation - methodische und konzeptionelle Ansätze der Wirkungsmessung in der Entwicklungszusammenarbeit. *Zeitschrift für Evaluation (ZfEv)*, 8 (2), 183-213.
- Caspari, A. (2015). Well done? Who knows... Ein Plädoyer für Meta-Evaluationen. In V. Hennefeld, W. Meyer, S. Silvestrini & R. Stockmann (Hrsg.), *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft: Festschrift zum 60. Geburtstag von Reinhard Stockmann* (S. 143-165). Münster: Waxmann.
- Caspari, A. & Barbu, R. (2008). *Wirkungsevaluierungen: Zum Stand der internationalen Diskussion und dessen Relevanz für Evaluierungen der deutschen Entwicklungszusammenarbeit*. (Evaluation Working Papers). Bonn: BMZ.
- Cassen, R. H. (1986). *Does aid work? Report to an intergovernmental task force*. Oxford: Clarendon Press.
- Catsambas, T. T., Segone, M., Silva, S. de & Saunders, M. (2013). EvalPartners: an international partnership to strengthen civil society's evaluation capacities and promote equity. In M. Segone & J. Rugh (Hrsg.), *Evaluation and Civil Society Stakeholders' perspectives on National Evaluation Capacity Development* (43-50). New York: UNICEF.
- Chambers, R. (1981). Rapid rural appraisal. Rationale and repertoire. *Public Administration and Development*, 1 (2), 95-106. doi: 10.1002/pad.4230010202.
- Chambers, R. (1994). Participatory rural appraisal (PRA): Challenges, potentials and paradigm. *World Development*, 22 (7), 953-969.
- Chawansky, M., Hayhurst, L., McDonald, M. & van Ingen, C. (2017). Innovations in sport for development and peace research. *Third World Thematics: A TWQ Journal*, 2 (1), 1-6. doi: 10.1080/23802014.2017.1374208.
- Chilisa, B. (2015, 31. August). *A Synthesis Paper on the Made in Africa Evaluation (MAE) Concept* (African Evaluation Association (AfrEA), Hrsg.). Zugriff am 24. November 2019 unter <https://afrea.org/wp-content/uploads/2018/06/MAE2-Final-31st-august.pdf>
- Cloete, F., Rabie, B. & Coning, C. de (Hrsg.). (2014). *Evaluation management in South Africa and Africa* (First edition). Stellenbosch: SUN Media.
- Coakley, J. (2011). Youth Sports: What Counts as "Positive Development?" *Journal of Sport and Social Issues*, 35 (3), 306-324. doi: 10.1177/0193723511417311.
- Coalter, F. (2007a). *A wider social role for Sport. Who's keeping the score?* New York: Routledge.
- Coalter, F. (2007b). Sports Clubs, Social Capital and Social Regeneration: 'ill-defined interventions with hard to follow outcomes'? *Sport in Society*, 10 (4), 537-559. doi: 10.1080/17430430701388723.
- Coalter, F. (2006). *Sport-in-Development. A Monitoring and Evaluation Manual*. University of Stirling. Zugriff am 20. Januar 2020 unter <https://www.sportanddev.org/en/article/publication/sport-development-monitoring-and-evaluation-manual>

- Coalter, F. (2009). Sport-in Development: Theoretical Frameworks. In R. Levermore & A. Beacom (Hrsg.), *Sport and international development* (S. 55-75). Basingstoke: Palgrave Macmillan.
- Coalter, F. (2010). The politics of sport-for-development: Limited focus programmes and broad gauge problems? *International Review for the Sociology of Sport*, 45 (3), 295-314. doi: 10.1177/1012690210366791.
- Coalter, F. (2013a). *Sport for development. What game are we playing?* New York: Routledge.
- Coalter, F. (2013b). 'There is loads of relationships here': Developing a programme theory for sport-for-change programmes. *International Review for the Sociology of Sport*, 48 (5), 594-612. doi: 10.1177/1012690212446143.
- Coalter, F. & Taylor, J. (2010). *Sport-for-development impact study*: University of Stirling. Zugriff am 21. April 2020 unter <https://www.idrettsforbundet.no/contentassets/321d9aedf8f64736ae55686d14ab79bf/fredcoaltersseminalmandemannual1.pdf>
- Cohen, A. & Welty Peachey, J. (2015). The making of a social entrepreneur: From participant to cause champion within a sport-for-development context. *Sport Management Review*, 18 (1), 111-125. doi: 10.1016/j.smr.2014.04.002.
- Collison, H. (2019). SDP and social anthropology. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 218-229). Milton: Routledge.
- Collison, H., Darnell, S. C., Giulianotti, R. & Howe, P. D. (Eds.). (2019). *Routledge Handbook of Sport for Development and Peace*. Milton: Routledge.
- Collison, H., Giulianotti, R., Howe, P. D. & Darnell, S. (2016). The methodological dance: critical reflections on conducting a cross-cultural comparative research project on 'Sport for Development and Peace'. *Qualitative Research in Sport, Exercise and Health*, 8 (5), 413-423. doi: 10.1080/2159676X.2016.1206610.
- Collison, H. & Marchesseault, D. (2018). Finding the missing voices of Sport for Development and Peace (SDP): using a 'Participatory Social Interaction Research' methodology and anthropological perspectives within African developing countries. *Sport in Society*, 21 (2), 226-242. doi: 10.1080/17430437.2016.1179732.
- Commonwealth Secretariat. (Juni, 2018). *Measuring the Contribution of Sport to the Sustainable Development Goals. Discussion Paper for the UN Expert Group Meeting on "Strengthening the Global Framework for Leveraging Sport for Development and Peace"*. Zugriff am 20. Mai 2020 unter <https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2018/06/8.pdf>
- Commonwealth Secretariat. (2019). *Measuring the contribution of sport, physical education and physical activity to the Sustainable Development Goals. Toolkit and model indicators*. London: Commonwealth Secretariat. Zugriff am 19. April 2020 unter <https://thecommonwealth.org/sites/default/files/inline/Sport-SDGs-Indicator-Framework.pdf>

- Cracknell, B. E. (1988). Evaluating development assistance. A review of the literature. *Public Administration and Development*, 8 (1), 75-83. doi: 10.1002/pad.4230080107.
- Cracknell, B. E. (2002). *Evaluating development aid. Issues, problems and solutions* (3rd printing). London: Sage.
- Creswell, J. W. & Creswell, J. D. (2018). *Research design. Qualitative, quantitative, and mixed methods approaches* (Fifth edition). Los Angeles: Sage.
- Cronin, Ó. (2011). *Comic Relief Review. Mapping the research on the impact of Sport and Development interventions: Comic Relief*.
- Cummings, F. H. (1997). Logic Models, Logical Frameworks and Results-Based Management: Contrasts and Comparisons. *Canadian Journal of Development Studies/Revue canadienne d'études du développement*, 18 (sup1), 587-596. doi: 10.1080/02255189.1997.9669719.
- Currie-Alder, B., Kanbur, R., Malone, D. M. & Medhora, R. (2014). *International Development. Ideas, experiences and prospects*: Oxford University Press.
- DAAD. (Mai 2018). *Die Umsetzung der „Agenda 2030 für nachhaltige Entwicklung“ im Deutschen Akademischen Austauschdienst (DAAD). Stand und Herausforderungen*. DAAD Standpunkt. Zugriff am 19. Dezember 2021 unter https://www2.daad.de/medien/der-daad/unsere-mission/standpunkte/standpunkt_agenda2030.pdf
- Dabelstein, N. & Patton, M. Q. (2015). The Paris Declaration on Aid Effectiveness. History and Significance. *The Canadian Journal of Program Evaluation*, 27 (3), 19-36.
- Dahler-Larsen, P. (2017). Die sich verändernde Rolle der Evaluation in einer sich verändernden Gesellschaft. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 21-34). Münster: Waxmann Verlag GmbH.
- Damerius, I. (2008). *Sport und interkulturelle Verständigung. Konzeption einer qualitativen Beobachtungsstudie im Rahmen einer Interventionsmaßnahme in Israel*. Diplomarbeit, Deutsche Sporthochschule Köln. Köln.
- Darnell, S. C. (2007). Playing with Race: Right to Play and the Production of Whiteness in 'Development through Sport'. *Sport in Society*, 10 (4), 560-579. doi: 10.1080/17430430701388756.
- Darnell, S. C. (2012). *Sport for Development and Peace. A Critical Sociology*. London: Bloomsbury Academic.
- Darnell, S. C. (2014). Orientalism through sport: towards a Said-ian analysis of imperialism and 'Sport for Development and Peace'. *Sport in Society*, 17 (8), 1000-1014. doi: 10.1080/17430437.2013.838349.
- Darnell, S. C. (2019). SDP and the environment. In H. Collison, S. C. Darnell, R. Giulianiotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 396-405). Milton: Routledge.

- Darnell, S. C. & Black, D. R. (2011). Mainstreaming Sport into International Development Studies. *Third World Quarterly*, 32 (3), 367-378. doi: 10.1080/01436597.2011.573934.
- Darnell, S. C., Chawansky, M., Marchesseault, D., Holmes, M. & Hayhurst, L. (2018a). The State of Play: Critical sociological insights into recent 'Sport for Development and Peace' research. *International Review for the Sociology of Sport*, 53 (2), 133-151. doi: 10.1177/1012690216646762.
- Darnell, S. C. & Dao, M. (2017). Considering sport for development and peace through the capabilities approach. *Third World Thematics: A TWQ Journal*, 2 (1), 23-36. doi: 10.1080/23802014.2017.1314772.
- Darnell, S. C., Giulianotti, R., David Howe, P. & Collison, H. (2018b). Re-Assembling Sport for Development and Peace through Actor Network Theory: Insights from Kingston, Jamaica. *Sociology of Sport Journal*, 35 (2), 89-97. doi: 10.1123/ssj.2016-0159.
- Darnell, S. C. & Hayhurst, L. (2012). Hegemony, postcolonialism and sport-for-development: a response to Lindsey and Grattan. *International Journal of Sport Policy and Politics*, 4 (1), 111-124. doi: 10.1080/19406940.2011.627363.
- Darnell, S. C. & Hayhurst, L. M. (2014). De-Colonising the Politics and Practice of Sport-for-Development: Critical Insights from Post-Colonial Feminist Theory and Methods. In N. Schulenkorf & D. Adair (Eds.), *Global Sport-for-Development. Critical Perspectives* (pp. 33-61). Basingstoke: Palgrave Macmillan.
- Darnell, S. C., Whitley, M. A., Camiré, M., Massey, W. V., Blom, L. C., Chawansky, M. et al. (2019). Systematic Reviews of Sport for Development Literature: Managerial and Policy Implications. *Journal of Global Sport Management*, 1-18. doi: 10.1080/24704067.2019.1671776.
- Darnell, S. C., Whitley, M. A. & Massey, W. V. (2016). Changing methods and methods of change: reflections on qualitative research in Sport for Development and Peace. *Qualitative Research in Sport, Exercise and Health*, 8 (5), 571-577. doi: 10.1080/2159676X.2016.1214618.
- Davies, P. T. (2004). *Is evidence-based government possible? Jerry Lee lecture 2004*, Paper presented at the 4th annual Campbell Collaboration Colloquium, Washington DC, 19 February.
- Debiel, T. (Hrsg.). (2018). *Entwicklungspolitik in Zeiten der SDGs. Essays zum 80. Geburtstag von Franz Nuscheler*. Duisburg: Institut für Entwicklung und Frieden.
- Defila, R., Di Giulio, A. & Scheuermann, M. (2008). *Management von Forschungsbündeln - Möglichkeiten der Professionalisierung und Unterstützung* (Standpunkte (DFG)). Weinheim: Wiley-VCH.
- DeGEval. (2008). *Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren* (DeGEval - Gesellschaft für Evaluation, Hrsg.), Mainz.
- DeGEval. (2016). *Standards für Evaluation. Erste Revision 2016* (DeGEval - Gesellschaft für Evaluation, Hrsg.). Zugriff am 24. November 2019 unter https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf

- DeGEval. (2016a). *Checkliste zur Anwendung der Standards. Begleitmaterial zu den Standards für Evaluation, Gesellschaft für Evaluation e.V.* Zugriff am 22. Juni 2021 unter https://www.degeval.org/fileadmin/DeGEval-Standards/2018_Checkliste_zur_Anwendung_der_Standards.pdf
- DeGEval. (2018, 06. August). *Die Arbeitskreise der DeGEval.* Zugriff am 24. November 2019 unter <https://www.degeval.org/arbeitskreise/>
- DEval (2017). *Lernen fördern. Transparenz schaffen. 5 Jahre DEval 2012-2017.* Bonn: DEval.
- DEval. (2018). *Standards für Evaluierungen des DEval (Methoden und Standards 2018):* DEval. Zugriff am 13. Januar 2020 unter https://www.deval.org/files/content/Dateien/Evaluierung/Policy%20Briefs/DEval_Evaluierungsstandards_2018_web.pdf
- DEval (2019). *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit.* Bonn: DEval.
- DEval. (2019a). *Über uns. Handlungsfelder.* Zugriff am 06. Dezember 2019 unter <https://www.deval.org/de/ueber-uns.html>
- DEval. (2019, 17. Dezember). *„Kohärenz“ neues OECD-Evaluierungskriterium.* Zugriff am 13. Januar 2020 unter <https://www.deval.org/de/presse-meldung-detail/koh%C3%A4renz-neues-oecd-evaluierungskriterium.html>
- DEZA. (2005). *Sport für Entwicklung und Frieden.* (Direktion für Entwicklung und Zusammenarbeit (DEZA), Hrsg.). Zürich, Bern: Waldburger Consulting.
- DFB. (2016, 26. Mai). *Internationale Zusammenarbeit: DFB und BMZ/GIZ veranstalten Workshop.* Zugriff am 25. Oktober 2020 unter <https://www.dfb.de/internationales/news-detail/internationale-zusammenarbeit-dfb-und-bmzgiz-veranstalten-workshop-146824/>
- DFB. (2016, 29. Mai). *DFB und BMZ verstärken Zusammenarbeit.* Zugriff am 25. Oktober 2020 unter <https://www.dfb.de/internationales/news-detail/dfb-und-bmz-verstaerken-zusammenarbeit-147048/>
- DFB. (2019, 23. August). *Sport für Entwicklung: Fußball in Gaza, Kenia und Kolumbien.* Zugriff am 25. Oktober 2020 unter <https://www.dfb.de/news/detail/internationaler-kurs-sport-fuer-entwicklung-206428/>
- Dieckert, J. & Wopp, C. (Hrsg.). (2002). *Handbuch Freizeitsport* (Beiträge zur Lehre und Forschung im Sport, Bd. 134). Schorndorf: Hofmann.
- Dietrich, K., Jost, E., Schießl, M. & Smidt, T. (1986). *Traditionelle Bewegungskultur und moderner Sport: Ansätze für die Entwicklungszusammenarbeit.* Ahrensburg: Czwalina.
- Digel, H. (1978). Zusammenfassende Stellungnahme zum Fragenkatalog II, 1-11 (Sachverständige aus der Sportwissenschaft). In Dt. Bundestag (Hrsg.), *Stenografisches Protokoll über die Anhörung von Sachverständigen in der 20. Sitzung des Sportausschusses am Mittwoch, dem 29. November 1978.* Bonn: Bundeshaus.

- Digel, H. (1985). Zur Rolle des Sports im Prozeß der Entwicklung unentwickelter Gesellschaften. In R. Dohrmann, M. John & W. Küper (Hrsg.), *Sportförderung in Ländern der Dritten Welt. Aus der 10-jährigen Arbeit der GTZ in diesem Bereich* (Sonderpublikation der GTZ, S. 61-74). Roßdorf: TZ-Verlagsges.
- Digel, H. (1989). Entwicklungszusammenarbeit im Sport. Fremdes verstehen, voneinander lernen. In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 203-216). Schorndorf: Hofmann.
- Digel, H. & Fornoff, P. (1989). *Sport in der Entwicklungszusammenarbeit* (Forschungsberichte des Bundesministeriums für Wirtschaftliche Zusammenarbeit, Bd. 96). München: Weltforum Verlag.
- Dohrmann, R. (Hrsg.). (1987). *Sportförderung in Ländern der Dritten Welt. 10. ADL-Kongress, Kiel 1986* (Sportreport, Bd. 13). Arbeitsgruppe "Sportunterricht in Entwicklungsländern". Eschborn: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ).
- Dohrmann, R. (Hrsg.). (1988). *Sportförderung in Ländern der Dritten Welt, The Promotion of Sports in Third World Countries. Sports for All* (Sportreport, Bd. 17). Eschborn: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ).
- Dohrmann, R., John, M. & Küper, W. (Hrsg.). (1985). *Sportförderung in Ländern der Dritten Welt. Aus der 10-jährigen Arbeit der GTZ in diesem Bereich* (Sonderpublikation der GTZ). Roßdorf: TZ-Verlagsges.
- Doll-Tepper, G., Petry, K., Dumon, D., Koenen, K., Weinberg, B. & Meier, Marianne, Giselbrecht, Lena. (2013). *Sport und Entwicklung. Eine Übersicht über die bedeutendsten in Deutschland ansässigen Akteure und ihrer Aktivitäten*. Interner Bericht: i.A. der GIZ unter der Leitung des ICSSPE.
- Doll-Tepper, G., Talbot, M., Petry, K., Dumon, D., Weinberg, B., Koenen, K. et al. (2014). *Sektorvorhaben "Sport für Entwicklung". Eine Übersicht über bedeutende internationale Akteure und ihre Aktivitäten*. Interner Bericht: i.A. der GIZ unter der Leitung des ICSSPE; DSHS Köln.
- Donaldson, S. I. & Lipsey, M. W. (2007). Roles of Theory in Contemporary Evaluation Practice: Developing Practical Knowledge. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of evaluation. Policies, programs and practices* (pp. 56-75). London: Sage Publ.
- Donnelly, P., Atkinson, M., Boyle, S. & Szto, C. (2011). Sport for Development and Peace: a public sociology perspective. *Third World Quarterly*, 32 (3), 589-601. doi: 10.1080/01436597.2011.573947.
- Dorfman, R. (1963). *Measuring Benefits of Government Investments*. Washington D.C.: Brookings Institution.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (Springer-Lehrbuch, 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin: Springer.
- DOSB. (o.J.). *Internationale Sportförderung des Auswärtigen Amts. 50 Jahre Zusammenarbeit im Sport – Menschen bewegen, Grenzen überwinden*: Deutscher

- Olympischer Sportbund (DOSB). Zugriff am 23. Oktober 2020 unter https://www.dfb.de/fileadmin/_dfbdam/Int__Sportfoerder_DOSB_DEUTSCH_RZ.pdf
- Dt. Bundestag. (1970). *Sportförderung* (Verhandlungen des Deutschen Bundestages, Drucksachen 6/1122). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1973). *Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 7/1040). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1976). *3. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 7/4609). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1978). *4. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 8/2033). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1982). *5. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 9/1945). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1987). *6. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 10/6241). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1991). *7. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 11/8459). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1995). *8. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 13/1114). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (1999). *9. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 14/1859). Bonn: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (2002). *10. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 14/9517). Berlin: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (2006). *11. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 16/3750). Berlin: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (2010). *12. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 17/2880). Berlin: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dt. Bundestag. (2014). *13. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 18/3523). Berlin: Dtl., Bundesregierung, Presse- u. Informationsamt.

- Dt. Bundestag. (2019). *14. Sportbericht der Bundesregierung* (Verhandlungen des Deutschen Bundestages, Drucksachen 19/9150). Berlin: Dtl., Bundesregierung, Presse- u. Informationsamt.
- Dudfield, O. (2019). SDP and the Sustainable Development Goals. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 116-127). Milton: Routledge.
- Dudfield, O. & Dingwall-Smith, M. (2015). *Sport for Development and Peace and the 2030 Agenda for Sustainable Development*: Commonwealth Secretariat. Zugriff am 20. Januar 2020 unter https://thecommonwealth.org/sites/default/files/inline/CW_SDP_2030+Agenda.pdf
- Duflo, E., Glennerster, R. & Kremer, M. (2007). *Using Randomization in Development Economics Research: A Toolkit* (Discussion Paper Series Nr. 6059): Centre for Economic Policy Research (CEPR). Zugriff am 05. Dezember 2019 unter https://scholar.harvard.edu/files/kremer/files/randomization_toolkit_dev_economics.pdf
- Eberspächer, H. & Egger, K. (1978). Orientierungs- und Entscheidungshilfen für die Sportförderung in Entwicklungsländern. In Dt. Bundestag (Hrsg.), *Stenografisches Protokoll über die Anhörung von Sachverständigen in der 20. Sitzung des Sportausschusses am Mittwoch, dem 29. November 1978*. Bonn: Bundeshaus.
- Eckert, A. (2015). Geschichte der Entwicklungszusammenarbeit. *Aus Politik und Zeitgeschichte (APuZ)*, 65 (7-9), 3-8.
- Edwards, M. B. (2015). The role of sport in community capacity building: An examination of sport for development research and practice. *Sport Management Review*, 18 (1), 6-19. doi: 10.1016/j.smr.2013.08.008.
- Eichberg, H. (1973). Verhaltensumformung nach europäischem Vorbild? *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 93, 641-670.
- Elsemann, K., Hebel, M., Jäger, L. & Daraspe, C. (o.J.). *Monitoring and Evaluation in Sport for Development*: streetfootballworld. Zugriff am 19. April 2020 unter <https://www.streetfootballworld.org/sites/default/files/Impact%20Measurement%20Manual.pdf>
- Engagement Global. (2019). *Das Jahr 2019 auf einen Blick*. Zugriff am 02. November 2020 unter https://www.engagement-global.de/files/2_Mediathek/Mediathek_EG/Weitere_Publicationen/Engagement_ermoeglichen_Einleger_2019_bf_final.pdf
- Engelhardt, J. (2019). SDP, monitoring, and evaluation. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 128-140). Milton: Routledge.
- EU. (2014, 14. August). *Evaluation matters. The Evaluation Policy for European Union Development Co-operation*. Zugriff am 12. Dezember 2019 unter https://ec.europa.eu/europeaid/sites/devco/files/evaluation-matters_en_0.pdf
- Euler, M., Noltze, M. & Verspohl, I. (2018). *Evaluierung von Nachhaltigkeit in Zeiten der Agenda 2030*. DEval Policy Brief 4/2018: DEval.

- Eurostat. (2018). *Sustainable development in the European Union. Monitoring report on progress towards the SDGs in an EU context: 2018 edition* (2018 edition). Luxembourg: Publications Office of the European Union.
- Eurostat. (2019). *Sustainable development in the European Union. Overview of progress towards the SDGs in an EU context: 2019 edition* (Eurostat compact guides, 2019 edition). Luxembourg: Publications Office of the European Union.
- EvalPartners. (2015). *Evaluation of EvalPartners. The Global Movement to Strengthen National Evaluation Capacities*. Zugriff am 24. November 2019 unter <https://www.evalpartners.org/sites/default/files/documents/reports/EvalPartners-Evaluation-Full-Report-2015-02-03.pdf>
- EvalPartners. (2016, 15. Januar). *EvalAgenda2020. Global Evaluation Agenda 2016-2020*. Zugriff am 20. Januar 2020 unter <https://www.evalpartners.org/sites/default/files/documents/EvalAgenda2020.pdf>
- EvalSDGs. (o.D.). *Concept Note*. Zugriff am 15. Januar 2020 unter <https://evalsdgs.org/wp-content/uploads/2016/09/evalsdg-1-pager-back-to-back-sept14.pdf>
- EvalSDGs. (2021). *Welcome to EvalSDGs!* Zugriff am 18. Februar 2022 unter <https://evalsdgs.org/>
- Farello, A., Blom, L., Mulvihill, Thalia & Erickson, J. (2019). Understanding female youth refugees' experiences in sport and physical education through the self-determination theory. *Journal of Sport for Development*, 7 (13), 55-72.
- Faust, J. (2010). Wirkungsevaluierung in der Entwicklungszusammenarbeit. *Aus Politik und Zeitgeschichte (APuZ)*, 10, 41-46.
- Faust, J. (2019). Vorwort. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 7). Bonn: DEval.
- Faust, J. & Leiderer, S. (2010). Die Effektivität der Entwicklungszusammenarbeit - Ergebnisse des ökonomischen Ländervergleichs. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 166-188). Baden-Baden: Nomos mbH & Co. KG.
- Faust, J. & Messner, D. (2010). Ordnungspolitische Herausforderungen für eine wirksamere EZ. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 259-288). Baden-Baden: Nomos.
- Faust, J. & Neubert, S. (2010). Entwicklungspolitik zwischen Fundamentalkritik und Radikaloptimismus. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 15-26). Baden-Baden: Nomos.
- Faust, J. & Neubert, S. (Hrsg.). (2010). *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage). Baden-Baden: Nomos mbH & Co. KG.
- Faust, J. & Verspohl, I. (2018). *Die DAC-Evaluierungskriterien: Zwischen Optimierung und Transformation*. DEval Policy Brief 9/2018: DEval.
- Feinstein, O. (2019). Dynamic evaluation for transformational change. In van den Berg, R. D., C. Magro & S. Salinas Mulder (Hrsg.), *Evaluation for Transformational*

- Change: Opportunities and Challenges for the Sustainable Development Goals* (S. 17-32). Exeter, UK: IDEAS.
- Fiege, K. (2012). Planung und Durchführung von Gutachten: "Aktions- und entscheidungsorientierte Untersuchungen". In K. Fiege & T. Rauch (Hrsg.), *Entwicklungszusammenarbeit gestalten. Inhalte und Methoden für ein erfolgreiches Wirken in einem komplexen Berufsfeld* (Entwicklungspolitische Themenreihe des SLE, 1, S. 224-242). Berlin: Humboldt-Universität zu Berlin, Seminar für Ländliche Entwicklung (SLE).
- Fischer, K., Hauck, G. & Boatcă, M. (Hrsg.). (2016). *Handbuch Entwicklungsforschung* (Springer Nachschlagewissen). Wiesbaden: Springer VS.
- Fischer, K., Latiko, H. & Parthey, H. (Hrsg.). (2011). *Interdisziplinarität und Institutionalisierung der Wissenschaft* (Wissenschaftsforschung, Bd. 2010). Berlin: Wissenschaftlicher Verlag.
- Fischer, T. (2003). Pädagogik der Freizeit und Erlebnispädagogik. In R. Popp & M. Schwab (Hrsg.), *Pädagogik der Freizeit* (Basiswissen Pädagogik, Pädagogische Arbeitsfelder, 6, S. 177-194). Baltmannsweiler: Schneider Verlag Hohengehren.
- Flatters, B. & Pütz, H. (2020). Evaluation und Nachhaltigkeit. Bericht zur 22. Jahrestagung der DeGEval. *Zeitschrift für Evaluation (ZfEv)*, 19 (1), 186-190.
- Flick, U., Kardorff, E. v., Steinke, I. (Hrsg.). (2014). *Qualitative Forschung. Ein Handbuch*. Reinbek: Rowohlt.
- Florian, M., Krisch, F., Till, T. & Herrmanns, S. (2019). *Rigorous Impact Evaluation. Cross-section analysis – Main Report*. Bonn: GIZ GmbH.
- Foster, M., Brown, A. & Conway, T. (2000). *Sector-wide approaches for health development: a review of experience* (World Health Organization (WHO), Hrsg.), Genf. Zugriff am 14. Juni 2019 unter https://apps.who.int/iris/bitstream/handle/10665/66468/WHO_GPE_00.1.pdf?sequence=1&isAllowed=y
- Freeman, H. E., Rossi, P.H. & Wright, S. (1979). *Evaluating Social Projects in Developing Countries*. Paris: OECD.
- Frueh, S. (2019). Eine neue Ära für Evaluierungen. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 28-29). Bonn: DEval.
- Fues, T. (2010). Zur Wirksamkeit der Entwicklungszusammenarbeit der Vereinten Nationen. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 403-430). Baden-Baden: Nomos.
- Furubo, J.-E. (2017). Evaluation - ihr Erbe und ihre Zukunft. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 99-110). Münster: Waxmann Verlag GmbH.
- Furubo, J.-E., Rist, R. C. & Sandahl, R. (Eds.). (2002). *International atlas of evaluation* (Comparative policy analysis series). New Brunswick: Transaction Publ.
- Gadai, T. (2019). Sport for Development and Peace: Current Perspectives of Research. In D. Almeida Marinho, H. P. Neiva, C. P. Johnson & N. Mohamudally (Eds.), *Sports Science and Human Health. Different Approaches*. IntechOpen.

- Gaerte, S. (2011). *Fußball als Medium der bewegungsorientierten Gewaltprävention und gesellschaftlichen Entwicklung. eine empirische Evaluation der Arbeit von Colombianitos in Puerto Tejada - Kolumbien*. Diplomarbeit, Deutsche Sporthochschule Köln. Köln.
- Gall, H. (1989). Sportförderung als Bildungszusammenarbeit in Lateinamerika. In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 203-216). Schorndorf: Hofmann.
- Geuder-Jilg, E., Stadtmüller, H. & Harten, S. (2019). Evaluierungskapazitäten fördern – für mehr Eigenverantwortung und Rechenschaft. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 76-78). Bonn: DEval.
- Giebenhain, H. (1990). *Sportförderung in der Dritten Welt: eine Untersuchung über die Zusammenarbeit im Sport mit Entwicklungsländern durch die Bundesrepublik Deutschland im Lichte von Modernisierungs- und Dependenztheorien*. Dissertation, Eberhard-Karls-Universität Tübingen. Tübingen.
- Gilbert, K. & Bennett, W. (Eds.). (2012). *Sport, peace and development*. Champaign, Illinois: Common Ground Pub. LLC.
- Giulianotti, R. & Armstrong, G. (2014). The Sport for Development and Peace Sector: A Critical Sociological Analysis. In N. Schulenkorf & D. Adair (Eds.), *Global Sport-for-Development. Critical Perspectives* (pp. 15-32). Basingstoke: Palgrave Macmillan.
- Giulianotti, R., Coalter, F., Collison, H. & Darnell, S. C. (2019). Rethinking Sportland: A New Research Agenda for the Sport for Development and Peace Sector. *Journal of Sport and Social Issues*, 43 (6), 411-437. doi: 10.1177/0193723519867590.
- Giulianotti, R., Hognestad, H. & Spaaij, R. (2016). Sport for Development and Peace: Power, Politics, and Patronage. *Journal of Global Sport Management*, 1 (3-4), 129-141. doi: 10.1080/24704067.2016.1231926.
- GIZ. (o.D.a). *Fachexpertise: Monitoring*. Zugriff am 08. Juli 2020 unter <https://www.giz.de/fachexpertise/html/59828.html>
- GIZ. (o.D.b). *Fachexpertise: Externe Beratung*. Zugriff am 10. Oktober 2021 unter <https://www.giz.de/fachexpertise/html/61025.html>
- GIZ. (2013). *Unterstützung von Capacity Development. Ein Orientierungsrahmen für die Praxis*. Bonn, Eschborn: GIZ GmbH.
- GIZ. (2015a). *Kooperationsmanagement in der Praxis. Gesellschaftliche Veränderungen gestalten mit Capacity WORKS*. Wiesbaden: Springer Gabler.
- GIZ. (2015b). *Orientierung Scaling-up für Auftragsverantwortliche und Fachplaner/-innen*. Internes Dokument: GIZ.
- GIZ. (2017a). *Multi-Akteurs-Partnerschaften*. Eschborn: GIZ GmbH. Zugriff am 18. November 2019 unter <https://www.giz.de/en/downloads/giz2017-de-leistungsangebot-multiakteurspartnerschaften-1.pdf>
- GIZ. (2017). *Evaluierungsbericht 2017. Wissen, was wirkt*. Zugriff am 13. März 2020 unter <https://www.giz.de/en/downloads/giz2017-de-evaluierungsbericht.pdf>

- GIZ. (2018). *Projektdaten*, GIZ. Zugriff am 22. Mai 2018 unter https://www.giz.de/projektdaten/index.action?request_locale=de_DE
- GIZ. (Juli 2018). *Allgemeine Vertragsbedingungen (AVB) für die Erbringung von Dienst- und Werkleistungen im Auftrag der Deutschen GIZ GmbH*. Zugriff am 02. Februar 2022 unter <https://www.giz.de/de/downloads/giz2018-de-Allgemeine-Vertragsbedingungen-f%C3%BCr-die-Erbringung-von-Dienst-und-Werkleistungen-2018.pdf>
- GIZ (2018, 02. Juli). *Definition Sektorvorhaben* (E-Mail).
- GIZ. (2020, 31. Mai). *Kurzmeldungen im Mai 2020: Schnelle und lokale Lösungen gegen die Coronavirus-Pandemie. Vamos! Kinder in Lateinamerika fit halten – digitale Sportstunde mit ALBA Berlin*. Zugriff am 18. Januar 2022 unter <https://www.giz.de/de/mediathek/87056.html>
- GIZ. (September 2021). *Compliance Jahresbericht 2020*. Zugriff am 02. Februar 2022 unter <https://www.giz.de/de/downloads/giz2020-de-giz-compliance-jahresbericht.pdf>
- GIZ & DSHS. (2020). *S4D Impact Indicator Framework SDG 3*. Zugriff am 19. September 2021 unter <https://www.sport-for-development.com/imglib/downloads/Kompetenzen%20&%20IIFs/giz-dshs2020-en-sdg3-impact-indicator-framework.pdf>
- GIZ RV S4DA. (Mai 2017). *Regionalvorhaben „Sport für Entwicklung in Afrika“ (S4DA). Mehr Platz für Sport - Kein Platz ohne Konzept!* GIZ GmbH, Bonn.
- GIZ RV S4DA. (Mai 2017a). *Förderung beruflicher Bildung durch Sport. Sport für Entwicklung in Äthiopien*: GIZ GmbH, Bonn.
- GIZ RV S4DA. (Mai 2017b). *Gewaltprävention durch Sport im Fluchtkontext. Sport für Entwicklung in Kenia*: GIZ GmbH, Bonn.
- GIZ RV S4DA. (Mai 2017c). *Stärkung von Kommunen und Demokratieförderung durch Sport. Sport für Entwicklung in Mosambik*: GIZ GmbH, Bonn.
- GIZ RV S4DA. (Mai 2017d). *Förderung reproduktiver Gesundheit und Bildung durch Sport. Sport für Entwicklung in Namibia*: GIZ GmbH, Bonn.
- GIZ RV S4DA. (Mai 2017e). *Stärkung der Zivilgesellschaft und Förderung der sozialen und politischen Teilhabe durch Sport. Sport für Entwicklung in Togo*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (2018a). *Das Evaluierungssystem der GIZ. Allgemeine Beschreibung*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (2018b). *Das Evaluierungssystem der GIZ. Zentrale Projektevaluierung im BMZ-Geschäft*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (2018c). *Das Evaluierungssystem der GIZ. Unternehmensstrategische Evaluierung*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (2018d). *Das Evaluierungssystem der GIZ. Theorie des Wandels für Evaluierungen der GIZ*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (2018e). *Die Evaluierungspolicy der GIZ. Prinzipien, Leitlinien und Anforderungen an unsere Evaluierungspraxis*: GIZ GmbH, Bonn.
- GIZ Stabsstelle Evaluierung. (Dezember 2020). *Evaluierungsbericht 2020. Wissen nutzen*: GIZ GmbH, Bonn.

- GIZ SV SfE. (November 2017a). *Afghanistan: Förderung von Mädchen und jungen Frauen durch Sport*: GIZ GmbH, Bonn.
- GIZ SV SfE. (November 2017b). *Brasilien: Fußball, Gewaltprävention und Jugendförderung*: GIZ GmbH, Bonn.
- GIZ SV SfE. (November 2017c). *Namibia: Stärkung von Jugendlichen, insbesondere Mädchen und jungen Frauen durch Sport*: GIZ GmbH, Bonn.
- GIZ SV SfE. (November 2017d). *Palästinensische Gebiete: Förderung beruflicher Bildung durch Sport*: GIZ GmbH, Bonn.
- GIZ SV SfE. (Juli 2016). *Metodologia Treino Social 14+. Aprendendo no campo, Vencendo na vida*. Bonn, Eschborn: GIZ GmbH.
- GIZ SV SfE. (November 2016). *Mehr Platz für Sport - 1000 Chancen für Afrika* (BMZ, Hrsg.). GIZ GmbH, Bonn.
- GIZ SV SfE. (Oktober 2019). *Indonesia: Football as a tool for quality education in schools*: GIZ GmbH, Bonn.
- GIZ SV SfE. (März 2020). *Kolumbien: Friedensentwicklung und Gewaltprävention durch Sport*: GIZ GmbH, Bonn.
- GIZ SV SfE. (Oktober 2021). *1:0 für Bildung, Gesundheit und Integration. Wie Sport die Zukunftsperspektiven von Kindern und Jugendlichen verbessert*: GIZ GmbH, Bonn.
- GIZ SV SfE. (November 2022). *Zahlen Daten Fakten Master Tabelle*. Internes Dokument.
- GIZ/YDF. (2010, 29. Juni). *About us*. Zugriff am 18. Oktober 2020 unter http://www.zaydf.org/pages/who_we_are/
- Glöckner, S. (2010). *Soziale Prozesse im Sport und ihre pädagogischen Chancen, speziell in der Mannschaftssportart Fußball*. München: GRIN Verlag GmbH.
- GloKnoCo. (2019). *Global Sport for Development and Peace Knowledge Collaborative: Using sport as a tool for social change*. Zugriff am 29. Juni 2020 unter <https://gloknoco.net/>
- Goetze, D. (2002). *Entwicklungssoziologie. Eine Einführung* (Grundlagentexte Soziologie). Weinheim: Juventa-Verlag.
- Gohl, E. (2010). Participatory Impact Monitoring (PIM) / FAKT. In D. Oberndörfer, T. Hanf & H. Weiland (Hrsg.), *Verfahren der Wirkungsanalyse. Ein Handbuch für die entwicklungspolitische Praxis* (Freiburger Beiträge zu Entwicklung und Politik, 36, S. 120-127). Freiburg i. Br.: Arnold Bergstraesser Institut.
- Gomez, R. (2017). Die Bedeutung der DeGEval aus Sicht eines Auftraggebers von Evaluation: GIZ. *Zeitschrift für Evaluation (ZfEv)*, 16 (2), 157-164.
- GOPA. (o.D.). *Projects*. Zugriff am 18. Januar 2022 unter <https://www.gopa.de/projects?client=42&country=All®ion=25§or=All>
- Gothenburg Centre for Sustainable Development & University of Gothenburg. (o.D.). *SDG Impact Assessment Tool*. Zugriff am 12. März 2021 unter <https://sdgimpactassessmenttool.org/>
- GPEDC. (o.D.). *At a Glance*. Zugriff am 15. Januar 2020 unter <https://www.effectivecooperation.org/system/files/2021-01/English%20-infographic.pdf>

- Grafarend, K. & Gaerte, S. (2020). Entwicklung im Sport, des Sports und durch den Sport. Die internationale Arbeit des DOSB. In K. Petry (Hrsg.), *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage, S. 167-180). Leverkusen: Verlag Barbara Budrich.
- Grimm, S. (2010). Zur Wirksamkeit der europäischen Entwicklungspolitik. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 381-402). Baden-Baden: Nomos.
- Groll, M. (2005). *Transnationale Sportpolitik. Analyse und Steuerungsansatz sportpolitischer Interaktionen* (Sportforum, 14). Zugl.: Köln, Dt. Sporthochsch., Diss., 2005. Aachen: Meyer & Meyer.
- Groll, M. & Hillbring, M. (2011). Deutsche Entwicklungszusammenarbeit im Kontext sportpolitischer Handelns. In K. Petry, M. Groll & W. Tokarski (Hrsg.), *Sport und internationale Entwicklungszusammenarbeit. Theorie- und Praxisfelder* (S. 11-28). Köln: SportVerlag Strauß.
- Groll, T. (2019, 14. Oktober). *Armut verstehen, um sie zu bekämpfen*, ZEIT Online. Zugriff am 05. Dezember 2019 unter <https://www.zeit.de/wirtschaft/2019-10/nobelpreis-wirtschaft-abhit-banerjee-esther-duflo-michael-kremer-armutsforschung>
- Größing, S. (1993). *Bewegungskultur und Bewegungserziehung. Grundlagen einer sinnorientierten Bewegungspädagogik*. Schorndorf: Hofmann.
- Groves, L. C. & Hinton, R. B. (Eds.). (2004). *Inclusive aid. Changing power and relationships in international development* (Repr). London: Earthscan.
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, Calif: Sage.
- Guenther, D. E. (2006). Wirkungsorientierte Evaluation der Deutschen Welthungerhilfe. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 145-160.
- Guest, A. M. (2019). SDP and sports psychology. In H. Collison, S. C. Darnell, R. Giulianiotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 230-240). Milton: Routledge.
- Gutknecht-Gmeiner, M. (2014). Wie weit können, dürfen und müssen Evaluierende gehen, um die Welt zu retten? Zum Rollenverständnis von Evaluierenden. Frühjahrstagung des Arbeitskreises Aus- und Weiterbildung in der Evaluation am 21. März 2014 in Saarbrücken. *Zeitschrift für Evaluation (ZfEv)*, 13 (2), 364-371.
- Hageboek, M. (1984). United States Agency for International Development (USAID). In B. E. Cracknell (Ed.), *The evaluation of aid projects and programmes. Proceedings of the conference organised by the Overseas Development Administration in the Institute of Development Studies, at the University of Sussex 7-8 April 1983* (pp. 102-116). London: Overseas Development Administration.
- Hancock, M., Lyras, A. & Ha, J. P. (2013). Sport for Development programmes for girls and women: a global assessment. *Journal of Sport for Development*, 1 (1), 15-24.

- Harris, K. (2018). Building sport for development practitioners' capacity for undertaking monitoring and evaluation – reflections on a training programme building capacity in realist evaluation. *International Journal of Sport Policy and Politics*, 10 (4), 795-814. doi: 10.1080/19406940.2018.1442870.
- Harris, K. & Adams, A. (2016). Power and discourse in the politics of evidence in sport for development. *Sport Management Review*, 19 (2), 97-106. doi: 10.1016/j.smr.2015.05.001.
- Harten, S. & Petrova, M. (2018). Die deutsche EZ braucht MEHR Wirkungsorientierung. *Zeitschrift für Evaluation (ZfEv)*, 17 (1), 145-152.
- Hartmann, C. & Vorwerk, K. (2019). *Länderportfolioreviews. Ein Analyseinstrument für die deutsche Entwicklungszusammenarbeit*. DEval Policy Brief 4/2019: DEval.
- Hartmann, D. & Kwauk, C. (2011). Sport and Development. *Journal of Sport and Social Issues*, 35 (3), 284-305. doi: 10.1177/0193723511416986.
- Hartmann-Tews, I. (1996). *Sport für alle!? Strukturwandel europäischer Sportsysteme im Vergleich: Bundesrepublik Deutschland, Frankreich, Großbritannien*. Schorn-dorf: Hofmann.
- Hatton, D. (2015). *Shaping the 'Sport for Development' Agenda post-2015* (inFocus Enterprises Ltd, Hrsg.). Zugriff am 13. Mai 2020 unter https://impactin-focus.com/?jet_download=5591
- Hauck, L. (2019). Die Entwicklungszusammenarbeit der Zukunft. Herausforderungen der Finanziellen Zusammenarbeit meistern - Möglichkeiten und Grenzen der Evaluierung. In H. Sangmeister & H. Wagner (Hrsg.), *Die Entwicklungszusammenarbeit der Zukunft* (S. 133-158). Baden-Baden: Nomos.
- Haudenhuyse, R., Theeboom, M. & Nols, Z. (2013). Sports-based interventions for socially vulnerable youth: Towards well-defined interventions with easy-to-follow outcomes? *International Review for the Sociology of Sport*, 48 (4), 471-484. doi: 10.1177/1012690212448002.
- Hayhurst, L. M. (2009). The power to shape policy: charting sport for development and peace policy discourses. *International Journal of Sport Policy and Politics*, 1 (2), 203-227. doi: 10.1080/19406940902950739.
- Hayhurst, L. M., Giles, A. R. & Radforth, W. M. (2015). 'I want to come here to prove them wrong': using a post-colonial feminist participatory action research (PFPAR) approach to studying sport, gender and development programmes for urban Indigenous young women. *Sport in Society*, 18 (8), 952-967. doi: 10.1080/17430437.2014.997585.
- Hayhurst, L. M., Kay, T. & Chawansky, M. (Hrsg.). (2016). *Beyond sport for development and peace. Transnational perspectives on theory, policy and practice*. London: Routledge.
- Heckmair, B. & Michl, W. (2008). *Erleben und Lernen. Einführung in die Erlebnispädagogik*. München: Reinhardt.
- Heinemann, K. (1985). Sport und Entwicklungshilfe in Ländern der Dritten Welt. *Sportwissenschaft*, 15 (3), 227-244.

- Heinze, F., Reiter, S., Berg-Lupper, U., Wach, K. & Riedle, S. (2019). Multiple Handlungsanforderungen an Evaluierende. *Zeitschrift für Evaluation (ZfEv)*, 18 (2), 261-283.
- Hemmer, H.-R. (2007). Das Evaluierungssystem in der EZ - Eine Einführung. In H.-R. Hemmer (Hrsg.), *Zur Wirksamkeitsdebatte in Entwicklungszusammenarbeit (EZ)* (S. 1-12). Erfurt: 10. Entwicklungsökonomisches Seminar vom 17.-20. Januar 2007. Zugriff am 20. Juni 2019 unter http://verein-eff.de/resources/tagungsband_2007.pdf
- Hemmer, H.-R. (2010). Wirksamkeit und Evaluierung der Entwicklungszusammenarbeit auf der Projektebene. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 189-228). Baden-Baden: Nomos.
- Henne, K. (2017). Indicator culture in sport for development and peace: a transnational analysis of governance networks. *Third World Thematics: A TWQ Journal*, 2 (1), 69-86. doi: 10.1080/23802014.2017.1353892.
- Hills, S., Gómez Velásquez, A. & Walker, M. (2018). Sport as an analogy to teach life skills and redefine moral values: A case study of the 'Seedbeds of Peace' sport-for-development programme in Medellín, Colombia. *Journal of Sport for Development*, 6 (10), 19-31.
- Hillyer, S. J., Zahorsky, M. & Munroe, Amanda, Moran, Sarah. (2011). *Sport & Peace: Mapping the Field*: Georgetown University, M.A. Programme in Conflict Resolution. Zugriff am 19. April 2020 unter www.sportanddev.org/sites/default/files/downloads/finalreportsportandpeacemappingthefield.pdf
- Hinz, M. (2001). *Konzeptionen zur Sportförderung in der Dritten Welt mit Schwerpunkt Südafrika*. Stuttgart: Ibidem-Verlag.
- Hirschman, A. O. *Development Projects observed*. Washington D.C.: Brookings Institution.
- Holt, N. L., Deal, C. J., Pankow, K., Pynn, S. R. & Jorgensen, H. (2019). SDP and positive youth development. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 341-351). Milton: Routledge.
- Holt, N. L., Neely, K. C., Slater, L. G., Camiré, M., Côté, J., Fraser-Thomas, J. et al. (2017). A grounded theory of positive youth development through sport based on results from a qualitative meta-study. *International review of sport and exercise psychology*, 10 (1), 1-49. doi: 10.1080/1750984X.2016.1180704.
- Holzappel, S. (2014). *Darstellung von Ergebnissen in der Entwicklungszusammenarbeit. Risiken und Grenzen* (Dt. Institut für Entwicklungspolitik (DIE), Hrsg.) (Analysen und Stellungnahmen 4/2014).
- Huinink, J. (1995). *Warum noch Familie? Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft*. Frankfurt a.M.: Campus Verlag.
- Huinink, J. & Mayer, K. U. (1995). Einleitung. In J. Huinink & K. U. Mayer (Hrsg.), *Kollektiv und Eigensinn. Lebensverläufe in der DDR und danach* (S. 7-24). Berlin: De Gruyter.

- IAAF. (2006). *IAAF Kids' Athletics, A Team Event for Children. A Practical Guide for Kids' Athletics Animators*: International Association of Athletics Federations (IAAF), (Hrsg.).
- InFocus. (o.D.). *Impact Measurement and Management (IMM) e-learning platform*. Zugriff am 23. März 2021 unter <https://impactinfocus.com/e-learning/>
- InFocus. (2014). *inFocus Quality Standards: A Detailed Guide* (Aqumen Sports Consulting, Laureus Sport for Good Foundation, streetfootballworld, Hrsg.). Zugriff am 19. April 2020 unter https://www.streetfootballworld.org/sites/default/files/in-focus%20Quality%20Standards_0.pdf
- InFocus. (2016). *7 Steps to Effectice Impact Measurement* (inFocus Enterprises Ltd, Hrsg.). Zugriff am 19. April 2020 unter [https://www.undp.org/content/dam/ukraine/docs/DG/socinnov/7%20Steps%20to%20Effective%20Impact%20Measurement_v3_13.12.16%20\(1\).pdf](https://www.undp.org/content/dam/ukraine/docs/DG/socinnov/7%20Steps%20to%20Effective%20Impact%20Measurement_v3_13.12.16%20(1).pdf)
- InFocus. (2021, 10. Juni). *Home - inFocus*. Zugriff am 24. März 2022 unter <https://impactinfocus.com/>
- IOCE. (2019). *International Organization for Cooperation in Evaluation - Growth*. Zugriff am 24. November 2019 unter <https://www.ioce.net/>
- Jacob, S., Speer, S. & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21 (1), 6-31. doi: 10.1177/1356389014564248.
- Jacobsen, D. (2015). *Die sportbezogene internationale Entwicklungszusammenarbeit von in Deutschland ansässigen Organisationen. Ausgewählte Akteure, Inhalte, Strategien*. Dissertation, Freie Universität Berlin. Berlin.
- Jahn, U. (2006). Evaluierungen in der Praxis der KfW Entwicklungsbank. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 95-124.
- Janus, H. & Klingebiel, S. (2016). *Ergebnisbasierte Ansätze: Ein Schlüssel für die entwicklungspolitische Umsetzung der 2030 Agenda?* (Dt. Institut für Entwicklungspolitik (DIE), Hrsg.) (Die aktuelle Kolumne vom 09.05.2016).
- Jeanes, R. & Lindsey, I. (2014). Where's the "Evidence"? Reflecting Monitoring and Evaluation within Sport-for-Development. In K. Young & C. Okada (Eds.), *Sport, Social Development and Peace* (Research in the Sociology of Sport, v.8, pp. 197-217). Bradford: Emerald Group Publishing Limited.
- John, M. (1989). Die Aufgabe der Deutschen Gesellschaft für Technische Zusammenarbeit als bundeseigene Durchführungsorganisation und ihre Maßnahmen zur Förderung des Sports in der Dritten Welt. In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 163-169). Schorndorf: Hofmann.
- Jones, G. J., Edwards, M. B., Bocarro, J. N., Bunds, K. S. & Smith, J. W. (2017). An integrative review of sport-based youth development literature. *Sport in Society*, 20 (1), 161-179. doi: 10.1080/17430437.2015.1124569.
- Jsfed.org. (o.D.). *About JSFD*. Zugriff am 23. Mai 2020 unter <https://jsfd.org/about-jsfd/>

- Kamla, B. (1985). *Participatory Evaluation: Are we on the right track?* Report of Secunderabad, Workshop on Participatory Self-Evaluation: Ideas and Action 2, Freedom from Hunger Campaign Rom: FAO, mimeo.
- Kantowsky, D. (Hrsg.). (1977). *Evaluierungsforschung und -praxis in der Entwicklungshilfe*. Zürich: Verlag d. Fachvereine a.d. schweizerischen Hochschulen u. Techniken.
- Kaplan, M., Vorwerk, K. & Leiderer, S. (2019). Von der Erklärung von Paris 2030: Überfordert die globale Nachhaltigkeitsagenda die Entwicklungszusammenarbeit? In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 18-19). Bonn: DEval.
- Karkoschka, O. (2017). *GIZ Wirkungsdaten 2016. Qualitative ergänzende Auswertung*. Bonn: GIZ GmbH.
- Karkoschka, O., Milcher, S., Müller-Gerold, J. & Bendtfeld-Huthmann, L. (2017). *Wirkungen global erfassen und ihren Kontext sehen. Herausforderungen aggregierter Berichterstattung am Beispiel der GIZ Wirkungsdaten*. Bonn: GIZ GmbH.
- Katenkamp, O. (2001). *Implizites Wissen in Organisationen: Konzepte, Methoden und Ansätze im Wissensmanagement*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kaufman, Z. A., DeCelles, J., Bhauti, K., Hershov, R. B., Weiss, H. A., Chaibva, C. et al. (2016). A Sport-Based Intervention to Increase Uptake of Voluntary Medical Male Circumcision Among Adolescent Male Students: Results from the MCUTS 2 Cluster-Randomized Trial in Bulawayo, Zimbabwe. *Journal of acquired immune deficiency syndromes (1999)*, 72 Suppl 4, S292-8. doi: 10.1097/QAI.0000000000001046.
- Kaufman, Z. A., Rosenbauer, B. P. & Moore, G. (2014). Lessons learned from Monitoring and Evaluating Sport-for-Development Programmes in the Caribbean. In N. Schulenkorf & D. Adair (Eds.), *Global Sport-for-Development. Critical Perspectives* (pp. 173-193). Basingstoke: Palgrave Macmillan.
- Kay, T. (2009). Developing through sport: evidencing sport impacts on young people. *Sport in Society*, 12 (9), 1177-1191. doi: 10.1080/17430430903137837.
- Kay, T. (2012). Accounting for legacy: monitoring and evaluation in sport in development relationships. *Sport in Society*, 15 (6), 888-904. doi: 10.1080/17430437.2012.708289.
- Kay, T. & Dudfield, O. (2013). *The Commonwealth Guide to Advancing Development through Sport*. London: Commonwealth Secretariat.
- Kay, T. & Mansfield, L. (2016). Researching with Go Sisters, Zambia: reciprocal learning in sport for development. In L. M. Hayhurst, T. Kay & M. Chawansky (Hrsg.), *Beyond sport for development and peace. Transnational perspectives on theory, policy and practice* (S. 214-229). London: Routledge.
- Kidd, B. (2008). A new social movement: Sport for development and peace. *Sport in Society*, 11 (4), 370-380.

- Kidd, B. (2011). Cautions, Questions and Opportunities in Sport for Development and Peace. *Third World Quarterly*, 32 (3), 603-609. doi: 10.1080/01436597.2011.573948.
- Kind, S. & Wessels, J. (2008). Das Spannungsfeld Evaluation und Begleitforschung - Komplexe Kompetenzerwartungen und Rollenkonflikte. *Zeitschrift für Evaluation (ZfEv)*, 7 (2), 347-352.
- Kirsch, A. (1978). Beantwortung des Fragenkataloges zur öffentlichen Anhörung über „Sportförderung in den Entwicklungsländern“. In Dt. Bundestag (Hrsg.), *Stenografisches Protokoll über die Anhörung von Sachverständigen in der 20. Sitzung des Sportausschusses am Mittwoch, dem 29. November 1978*. Bonn: Bundeshaus.
- Kjekshus, H. (1991). Evaluation Policy and Performance: The State of the Art. In O. Stokke (Ed.), *Evaluating development assistance. Policies and performance* (EADI book series, vol. 12, pp. 60-70). London: Cass.
- Klingebiel, S. (2013). *Entwicklungszusammenarbeit - eine Einführung* (Studies, 73). Bonn: Deutsches Institut für Entwicklungspolitik.
- Klingebiel, S. (2017). Entwicklungsforschung und entwicklungspolitische Kooperationsansätze. Von der Notwendigkeit eines neuen Designs. *Zeitschrift für Außen- und Sicherheitspolitik*, 10 (4), 453-468. doi: 10.1007/s12399-017-0662-9.
- Klingebiel, S. (2019). Rahmenbedingungen und Eckpunkte von Entwicklungszusammenarbeit in der Zukunft. In H. Sangmeister & H. Wagner (Hrsg.), *Die Entwicklungszusammenarbeit der Zukunft* (S. 41-54). Baden-Baden: Nomos.
- Klingebiel, S., Leiderer, S. & Schmidt, P. (2007). *Wie wirksam sind neue Modalitäten der Entwicklungszusammenarbeit? Erste Erfahrungen mit Programm-Based Approaches (PBAs)* (Discussion paper/DIE, 2007,7). Bonn: Dt. Institut für Entwicklungspolitik (DIE).
- Klingmann, D., Mialkas, C. & Weinberg, B. (2020). Die Rolle des Sports in der deutschen Entwicklungspolitik - Strategien, Strukturen und Prozesse. In K. Petry (Hrsg.), *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage, S. 143-154). Leverkusen: Verlag Barbara Budrich.
- Kloke-Lesch, A. & Messner, D. (2019). Weitermachen wie bisher ist keine Option. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 14-17). Bonn: DEval.
- Koch, H. (2021, 20. Oktober). *Gerd Müller hat geliefert*, taz.de. Zugriff am 20. November 2021 unter <https://taz.de/Scheidender-Entwicklungsminister/!5806100/>
- Kohnert, D. (1998). Lehren aus 15 Jahren ZOPP: Plädoyer für zielorientierte statt sektororientierte Steuerung der EZ. *E+Z Entwicklung und Zusammenarbeit* (Nr. 5/6), 137-140.
- König, J. (2011). Entwicklung und Nachhaltigkeit. Kritische Betrachtung von zwei dehnbaren Konzepten. In J. König & J. Thema (Hrsg.), *Nachhaltigkeit in der Entwicklungszusammenarbeit. Theoretische Konzepte, strukturelle Herausforderungen*

- und praktische Umsetzung* (Globale Gesellschaft und internationale Beziehungen, 1. Aufl., S. 41-68). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Koy, J. (2018). Einfluss von Wirkungszielen und Indikatoren auf PME, Organisation und Programme der EU. *Zeitschrift für Evaluation (ZfEv)*, 17 (1), 139-144.
- Krämer, M., Almqvist, O., Raetzell, L., Alber, B. & Verfürden, L. (2019). *Meta-Evaluierung und statistische Auswertung der Projektevaluierungen 2017 / 2018. Teil I: Meta-Evaluierung, im Auftrag der GIZ durchgeführt durch externe Evaluator/innen*. Bonn: GIZ GmbH.
- Kranz-Plote, J. (2006). Chancen und Herausforderungen bei der operativen Umsetzung der Millennium-Entwicklungsziele. Eine Innenperspektive. In F. Nuscheler & M. Roth (Hrsg.), *Die Millennium-Entwicklungsziele. Entwicklungspolitischer Königsweg oder ein Irrweg?* (Eine Welt, N.F., 20, S. 81-97). Bonn: Dietz.
- Krempin, M. (2019). Stärken ausbauen und neue Herausforderungen meistern: Die Zukunft der Technischen Zusammenarbeit der GIZ. In H. Sangmeister & H. Wagner (Hrsg.), *Die Entwicklungszusammenarbeit der Zukunft* (S. 119-132). Baden-Baden: Nomos.
- Kröger, C. (1989). Sport und Wissenschaft in der Kooperation am Beispiel des Internationalen Volleyball-Verbandes. In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 122-134). Schorndorf: Hofmann.
- Kromrey, H. (2001). Evaluation - ein vielschichtiges Konzept: Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis. *Sozialwissenschaften und Berufspraxis*, 24 (2), 105-131.
- Kromrey, H. (2003). Evaluation in Wissenschaft und Gesellschaft. *Zeitschrift für Evaluation (ZfEv)*, 2 (1), 93-116.
- Kromrey, H. (2007). Begleitforschung und Evaluation – fast das Gleiche, und doch etwas Anderes! In M. Glaser & S. Schuster (Hrsg.), *Evaluation präventiver Praxis gegen Rechtsextremismus. Positionen, Konzepte und Erfahrungen* (S. 113-135). Halle: Deutsches Jugendinstitut e. V.
- Kron, F. W., Jürgens, E. & Standop, J. (2013). *Grundwissen Pädagogik. Mit ... 12 Tabellen*. München: Reinhardt.
- Kroner, F. (2015). *The socio-psychological effects of a "Football-For-Development Program" pertaining to adolescent girls in Namibia. A qualitative analysis derived from the "Galz & Goals Program" aimed at improving the life skills of the participating girls*. Masterarbeit, Deutsche Sporthochschule Köln. Köln.
- Kubek, D., von der Ahe & B. (2015). *Mit Sport mehr bewegen! Report über wirkungsvolles zivilgesellschaftliches Engagement*. Berlin: Phineo.
- Kurz, B. & Kubek, D. (2018). *Kursbuch Wirkung. Das Praxishandbuch für alle, die Gutes noch besser tun wollen; mit Schritt-für-Schritt-Anleitungen & Beispielen* (5. Aufl.). Berlin: Phineo.
- Kurz, D. (2008). Der Auftrag des Schulsports. *Sportunterricht*, 57 (7), 211-218.
- Lachenmann, G. (1977). Evaluationsforschung - Historische Hintergründe, sozialpolitische Zusammenhänge und wissenschaftliche Einordnung. In D. Kantowsky

- (Hrsg.), *Evaluierungsforschung und -praxis in der Entwicklungshilfe* (S. 25-87). Zürich: Verlag d. Fachvereine a.d. schweizerischen Hochschulen u. Techniken.
- Langer, L. (2015). Sport for development – a systematic map of evidence from Africa. *South African Review of Sociology*, 46 (1), 66-86. doi: 10.1080/21528586.2014.989665.
- Lee, B. (2000). Theories of Evaluation. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 127-164). Opladen: Leske + Budrich.
- Leeuw, F. (2017). Cyber Society und die "Verdschungelung" von regeln. Neue Herausforderungen für die Evaluation im 21. Jahrhundert. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 141-155). Münster: Waxmann Verlag GmbH.
- Leeuw, F. & Vaessen, J. (2009). *Impact Evaluations and Development. Nonie Guidance on Impact Evaluation*. Washington D.C.: NONIE.
- Levermore, R. (2008). Sport: a new engine for development? *Progress in Development Studies*, 8 (2), 183-190. doi: 10.1177/146499340700800204.
- Levermore, R. (2009). Sport-in-International Development: Theoretical Frameworks. In R. Levermore & A. Beacom (Hrsg.), *Sport and international development* (S. 26-54). Basingstoke: Palgrave Macmillan.
- Levermore, R. (2011). Evaluating sport-for-development. *Progress in Development Studies*, 11 (4), 339-353. doi: 10.1177/146499341001100405.
- Levermore, R. & Beacom, A. (2009). Sport and Development: Mapping the Field. In R. Levermore & A. Beacom (Hrsg.), *Sport and international development* (S. 1-25). Basingstoke: Palgrave Macmillan.
- Levermore, R. & Beacom, A. (2012). Reassessing sport-for-development: moving beyond 'mapping the territory'. *International Journal of Sport Policy and Politics*, 4 (1), 125-137. doi: 10.1080/19406940.2011.627362.
- Levine, R. & Savedoff, B. (2007, 1. Dezember). *International Initiative for Impact Evaluation (3IE) – Beyond good intentions*. Zugriff am 16. Oktober 2019 unter <https://www.alliancemagazine.org/feature/international-initiative-for-impact-evaluation-3ie-beyond-good-intentions/>
- Lindsey, I. (2008). Conceptualising sustainability in sports development. *Leisure Studies*, 27 (3), 279-294. doi: 10.1080/02614360802048886.
- Lindsey, I. (2017). Governance in sport-for-development: Problems and possibilities of (not) learning from international development. *International Review for the Sociology of Sport*, 52 (7), 801-818. doi: 10.1177/1012690215623460.
- Lindsey, I. & Bitugu, B. B. (2019). Partnerships in and around SDP. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 80-90). Milton: Routledge.
- Lindsey, I. & Chapman, T. (2017). *Enhancing the contribution of sport to the sustainable development goals*. London: Commonwealth Secretariat.
- Lindsey, I., Chapman, T. & Dudfield, O. (2020). Configuring relationships between state and non-state actors: a new conceptual approach for sport and development.

- International Journal of Sport Policy and Politics*, 12 (1), 127-146. doi: 10.1080/19406940.2019.1676812.
- Lindsey, I. & Darby, P. (2019). Sport and the Sustainable Development Goals: Where is the policy coherence? *International Review for the Sociology of Sport*, 54 (7), 793-812. doi: 10.1177/1012690217752651.
- Lindsey, I. & Grattan, A. (2012). An 'international movement'? Decentring sport-for-development within Zambian communities. *International Journal of Sport Policy and Politics*, 4 (1), 91-110. doi: 10.1080/19406940.2011.627360.
- Lindsey, I., Kay, T., Jeanes, R. & Banda, D. (Eds.). (2017). *Localizing global sport for development* (Globalizing sport studies). Manchester: Manchester University Press.
- Lindsey, I., Kay, T., Jeanes, R. & Banda, D. (2017). Sport for development in policy, practice and research. In I. Lindsey, T. Kay, R. Jeanes & D. Banda (Eds.), *Localizing global sport for development* (Globalizing sport studies, pp. 17-51). Manchester: Manchester University Press.
- Loewe, M. (2010). Entwicklungspolitik, Armutsbekämpfung und Millennium Development Goals. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 101-138). Baden-Baden: Nomos.
- Löwen. (2001). *Eigenmaßnahmen im Kinder- und Jugendbereich - Innovationen und lessons learned aus der Praxis der TZ. Auswertung von Eigenmaßnahmen der Kinder- und Jugendförderung von 1994 bis 2001 Ansätze - Methoden - Wirkungen*. Eschborn: GTZ.
- Lücking, K., Bettighofer, S. & Freund, S. (2015a). *Erhebung zur Ausgangslage des DEval und seines Umfeldes* (DEval, Hrsg.), Bonn. Zugriff am 12. Dezember 2019 unter https://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2015_DEval_Baseline.pdf
- Lücking, K., Bettighofer, S. & Freund, S. (2016). "Systemprüfung" - Ein Impulsgeber für die Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit? DEval Policy Brief 02/2016: DEval.
- Lücking, K., Freund, S. & Bettighofer, S. (2015b). *Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit. Umsetzungsmonitoring der letzten Systemprüfung und Charakterisierung wesentlicher Elemente* (DEval, Hrsg.), Bonn. Zugriff am 12. Dezember 2019 unter https://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2015_DEval_Umsetzungsmonitoring.pdf
- Lüders, C. (2006). Qualitative Evaluationsforschung - was heißt hier Forschung? In U. Flick (Hrsg.), *Qualitative Evaluationsforschung. Konzepte - Methoden - Umsetzung* (S. 33-62). Reinbek: Rowohlt.
- Lyras, A. & Welty Peachey, J. (2011). Integrating sport-for-development theory and praxis. *Sport Management Review*, 14 (4), 311-326. doi: 10.1016/j.smr.2011.05.006.
- Lyras, A. & Welty Peachey, J. (2018). The conception, development, and application of sport-for-development theory. In G. B. Cunningham, J. S. Fink & A. Doherty

- (Eds.), *Routledge handbook of theory in sport management* (Routledge international handbooks, pp. 131-142). Abingdon: Routledge.
- MacArthur, J. D. (1995). *The evaluation of development projects. A review of the approaches and experiences of donor agencies*. Bradford: Univ.
- Magee, J. & Jeanes, R. (2013). Football's coming home: A critical evaluation of the Homeless World Cup as an intervention to combat social exclusion. *International Review for the Sociology of Sport*, 48 (1), 3-19. doi: 10.1177/1012690211428391.
- Magro, C. & van den Berg, R. D. (2019). Systems evaluations for transformational change. In van den Berg, R. D., C. Magro & S. Salinas Mulder (Hrsg.), *Evaluation for Transformational Change: Opportunities and Challenges for the Sustainable Development Goals* (S. 131-156). Exeter, UK: IDEAS.
- Maguire, J. A. (2013). *Reflections on process sociology and sport. Walking the line* (Sport in the global society). London: Routledge.
- Maihold, G. (2019). Entwicklungszusammenarbeit am Scheideweg: strategische Weichenstellung und institutioneller Reformbedarf. In H. Sangmeister & H. Wagner (Hrsg.), *Die Entwicklungszusammenarbeit der Zukunft* (S. 17-40). Baden-Baden: Nomos.
- Mandigo, J., Corlett, J. & Ticas, P. (2016). Examining the role of life skills developed through Salvadoran physical education programs on the prevention of youth violence. *Journal of Sport for Development*, 4 (7), 25-38.
- Mark, M. M. & Henry, G. T. (2007). Methods for Policy-making and Knowledge Development Evaluations. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of evaluation. Policies, programs and practices* (pp. 317-339). London: Sage Publ.
- Marshall, S. K. & Barry, P. (2015). Community Sport for Development: Perceptions from Practice in Southern Africa. *Journal of Sport Management*, 29 (1), 109-121. doi: 10.1123/JSM.2012-0301.
- Martinuzzi, A. & Meyer, W. (2017). Evaluation nachhaltiger Entwicklung in einer globalen Gesellschaft. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 71-98). Münster: Waxmann Verlag GmbH.
- Massey, W. V. & Whitley, M. A. (2019). SDP and research methods. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 175-184). Milton: Routledge.
- Massey, W. V., Whitley, M. A., Blom, L. & Gerstein, L. H. (2015). Sport for development and peace: a systems theory perspective on promoting sustainable change. *International Journal of Sport Management and Marketing*, 16 (1/2), 18-35. doi: 10.1504/IJSM.2015.074921.
- Mayer, K. U. (Hrsg.). (1990). *Lebensverläufe und sozialer Wandel* (Kölner Zeitschrift für Soziologie und Sozialpsychologie: Sonderhefte, Bd. 31). Opladen: Westdt. Verlag.

- Mayring, P., Kerlen, C., Sheikh, S., Hense, J. & Schwab, O. (2017). Zwanzig Jahre DeGEval – Historie, Entwicklungen, Herausforderungen. *Zeitschrift für Evaluation (ZfEv)*, 16 (2), 13-31.
- McDonald, M. G. (2019). SDP and feminist theory. In H. Collison, S. C. Darnell, R. Giuliannotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 197-207). Milton: Routledge.
- McKean, R. N. (1966). *Efficiency in Government through Systems Analysis. With emphasis on water resources development* (Publications in operations research, 3). New York: John Wiley.
- Merker, F. & Petry, K. (2014). *Sport- und Vereinsstrukturen in Maputo, Mosambik*. Interner Bericht: DSHS Köln.
- Meyer, W. (2020). Evaluation of Sustainable Development Goals between Ambition and Reality. How the Agenda 2030 Challenges the Evaluation Practice. *Zeitschrift für Evaluation (ZfEv)*, 2020 (02), 221-238. doi: 10.31244/zfe.2020.02.02.
- Meyer, W., Funk, E. & Nguyen, P. L. (2017). Partizipation und Bewertung: Weiterbestehende methodologische Herausforderungen. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 177-204). Münster: Waxmann Verlag GmbH.
- Meyer, W., Naidoo, I., D'Errico, S., Hofer, S., Bajwa, M., Tello Pérez, L. A. et al. (2018). *VNR reporting needs evaluation: a call for global guidance and national action* (IIED Briefing Nr. 8). Zugriff am 15. Januar 2020 unter <https://pubs.iied.org/17446IIED/>
- Meyer, W. & Stockmann, R. (2017). Schlussfolgerungen: Gemeinsame Perspektiven für die Institutionalisierung von Evaluation? In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 243-257). Münster: Waxmann Verlag GmbH.
- Michels, H., Petry, K. & Tokarski, W. (2003). Pädagogik der Freizeit und Sport. In R. Popp & M. Schwab (Hrsg.), *Pädagogik der Freizeit* (Basiswissen Pädagogik, Pädagogische Arbeitsfelder, 6, S. 152-176). Baltmannsweiler: Schneider Verlag Hohengehren.
- Michl, W. (2009). *Erlebnispädagogik* (UTB Profile, 3049). München: Reinhardt.
- Michl, W. & Seidel, H. (Hrsg.). (2018). *Handbuch Erlebnispädagogik*. München: Ernst Reinhardt Verlag.
- Ministerium für Schule und Weiterbildung des Landes NRW. (2014). *Rahmenvorgaben für den Schulsport in Nordrhein-Westfalen* (Ministerium für Schule und Weiterbildung des Landes NRW, Hrsg.) (Schule in NRW Heft 5020). Zugriff am 17. November 2020 unter https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SI/HS/sp/Rahmenvorgaben_Schulsport_Endfassung.pdf
- Mishan, E. J. (1971). *Cost-benefit Analysis*. London: Unwin University Books.

- Mitmannsgruber, P. (2005). *Einführung in die Bewegungs- und Sportpädagogik. Lehrbehelf zum Seminar*: Akademieverbund Pädagogische Hochschule in Krems. Zugriff am 18. November 2020 unter <http://mitmannsgruber.net/wp-content/uploads/2015/06/H-4-Lehrbehelf.pdf>
- Mohr, H.-W. (1977). *Bestimmungsgründe für die Verbreitung von neuen Technologien* (Beiträge zur Verhaltensforschung, 21). Zugl.: Köln, Univ., Diss., 1976. Berlin: Duncker & Humblot.
- Mosley, P. (1986). Aid-effectiveness. The Micro-Macro Paradox. *IDS Bulletin*, 17 (2), 22-27. doi: 10.1111/j.1759-5436.1986.mp17002004.x.
- Moustakas, L., Petry, K. & Bauer, K. (März 2022). *Sport for Development University Mapping Report: Activities, Challenges and Recommendations*. Interner Bericht: DSHS Köln.
- Mouton, J. (2014). Programme Evaluation Designs and Methods. In F. Cloete, B. Rabie & C. de Coning (Hrsg.), *Evaluation management in South Africa and Africa* (First edition, S. 165-203). Stellenbosch: SUN Media.
- Müller, C. E. (2017). Kausale Wirkungsevaluation zwischen methodischem Anspruch und empirischer Praxis. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 205-242). Münster: Waxmann Verlag GmbH.
- Müller, H. (2004). *Sportförderung im Prozess der Modernisierung am Beispiel der Leichtathletik*. Dissertation, Universität Tübingen.
- Müller, W. (2015). Von der angewandten Sozialforschung zur Sozialforschung der Verwendung. In V. Hennefeld, W. Meyer, S. Silvestrini & R. Stockmann (Hrsg.), *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft: Festschrift zum 60. Geburtstag von Reinhard Stockmann* (S. 15-40). Münster: Waxmann.
- Mundy, K. & Proulx, K. (2019). *Making Evaluation work for the Achievement of SDG4 Target 5: Equality and Inclusion in Education* (UNESCO Internal Oversight Service (IOS) Evaluation Office, Hrsg.).
- Muth, N. & Wirsich, N. (2019). *Yalla Nela'B. Guidebook for Active and Healthy Communities*. Bonn, Eschborn: GIZ GmbH.
- Naish, J. (2017). Sport for Development and Peace. Alignment, administration, power. In A. Bairner, J. Kelly & J. W. Lee (Eds.), *Routledge handbook of sport and politics* (pp. 297-312). London: Routledge.
- Neeten, M. (2020). *Facilitating social cohesion between Syrian refugees and Turkish host communities. A conceptual & methodological approach of sport-for-social cohesion activities*. Masterarbeit, Deutsche Sporthochschule Köln.
- Neubert, D. (2016). Entwicklungspolitik: Programme, Institutionen und Instrumente. In K. Fischer, G. Hauck & M. Boatcă (Hrsg.), *Handbuch Entwicklungsforschung* (Springer Nachschlagewissen, S. 359-373). Wiesbaden: Springer VS.
- Neubert, S. (2010). Zur Methodendiskussion und zur Analyse von Wirkungen. In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 139-165). Baden-Baden: Nomos.

- Neubert, S. & Müller, R. (2010). Method for Impact Assessment Method of Impact Assessment of Programmes and Projects (MAPP). In D. Oberndörfer, T. Hanf & H. Weiland (Hrsg.), *Verfahren der Wirkungsanalyse. Ein Handbuch für die entwicklungspolitische Praxis* (Freiburger Beiträge zu Entwicklung und Politik, 36, S. 88-96). Freiburg i. Br.: Arnold Bergstraesser Institut.
- Neubert, S., Mack, D. & Roxin, H. (2014). Vom Methodenstreit zum Methodenmix - Chronologie und Stand der Entwicklungsevaluierung in Deutschland. In Böttcher, W., Kerlen, C., Maats, P., Schwab, O. & Sheikh, S. (Hrsg.), *Evaluation in Deutschland und Österreich. Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval - Gesellschaft für Evaluation* (S. 61-72). Münster: Waxmann.
- Neuhaus, M., Barthel, M., Goeke, W., Joppien, C., Kather, C., Nitsche, S. et al. (2016). *Fördern und fordern II, Gemeinsames Lehren und Lernen im kompetenzorientierten Sportunterricht* (Bezirksregierung Arnsberg, Hrsg.), Arnsberg. Zugriff am 17. November 2020 unter https://www.schulentwicklung.nrw.de/lehrpläne/upload/klp_SII/sp/Gemeinsames_Lehren_und_Lernen_im_Sportunterricht_2016_Internetversion.pdf
- Nicholls, S., Giles, A. R. & Sethna, C. (2011). Perpetuating the 'lack of evidence' discourse in sport for development: Privileged voices, unheard stories and subjugated knowledge. *International Review for the Sociology of Sport*, 46 (3), 249-264. doi: 10.1177/1012690210378273.
- Nickolmann, S. (2006). Zur Durchführung von Evaluierung bei MISEREOR. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 125-144.
- Nike. (2012). *Designed to Move: A Physical Activity Action Agenda*. Zugriff am 13. Mai 2020 unter <https://www.sportsthinktank.com/uploads/designed-to-move-full-report-13.pdf>
- Nolte, C. & Schopper, S. (2017). Moderne Lehr- und Lernmethoden in der Evaluation. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 125-140). Münster: Waxmann Verlag GmbH.
- Nölting, S. (September 2018). *Sport für Entwicklung. In Jordanien und im Irak nutzen Flüchtlinge, Vertriebene und Einheimische Sportangebote für ein gesundes Leben und ein friedliches Miteinander in der Gemeinde.*: GIZ GmbH, Bonn.
- Noltze, M. (2019). Die Evaluierung von Nachhaltigkeit in der deutschen Entwicklungszusammenarbeit. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 32-33). Bonn: DEval.
- Noltze, M., Euler, M. & Verspohl, I. (2018a). *Nachhaltigkeit in der deutschen Entwicklungszusammenarbeit. Meta-Evaluierung* (DEval, Hrsg.), Bonn.
- Noltze, M., Euler, M. & Verspohl, I. (2018b). *Nachhaltigkeit in der deutschen Entwicklungszusammenarbeit. Evaluierungssynthese* (DEval, Hrsg.), Bonn.
- Noltze, M. & Harten, S. (2017). Neuer Wissensbedarf. *E+Z Entwicklung und Zusammenarbeit* (e-Paper no. 7/2017), 16. Zugriff am 28. Februar 2020 unter

- <https://www.dandc.eu/de/article/monitoring-und-evaluierung-als-schluesselfuer-den-erfolg-globaler-nachhaltigkeitspolitik>
- Noltze, M. & Hoffmann, B. (2019). Reform der Evaluierungskriterien auf Nachhaltigkeit. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 38). Bonn: DEval.
- Noltze, M. & Schwedersky, T. (2017). Kontributionsanalysen: Für Komplexes bestens geeignet. In DEval (Hrsg.), *Lernen fördern. Transparenz schaffen. 5 Jahre DEval 2012-2017* (S. 32-33). Bonn: DEval.
- Nowak, W. (2014). The Evolution of Development Assistance. *Journal of US-China Public Administrations*, 11 (5), 454-462.
- NSD. (2009). *How to Monitor and Evaluate 'Sport for Development' Projects* (Network for Sport and Development (NSD), Hrsg.). Copenhagen: SELA Advisory Group.
- Nuscheler, F. (2008). *Die umstrittene Wirksamkeit der EZ* (Institut für Entwicklung und Frieden (INEF), Universität Duisburg-Essen, Hrsg.) (INEF-Report Nr. 93), Duisburg. Zugriff am 12. März 2020 unter <https://pdfs.semanticscholar.org/1ba9/22081af2ca9c5751f0fa450e1efbb660b184.pdf>
- Nuscheler, F. & Roth, M. (2006). Die Millennium-Entwicklungsziele: ihr Potenzial und ihre Schwachstellen. Eine kritische Zusammenfassung. In F. Nuscheler & M. Roth (Hrsg.), *Die Millennium-Entwicklungsziele. Entwicklungspolitischer Königsweg oder ein Irrweg?* (Eine Welt, N.F., 20, S. 15-42). Bonn: Dietz.
- Oberndörfer, D., Hanf, T. & Weiland, H. (Hrsg.). (2010). *Verfahren der Wirkungsanalyse. Ein Handbuch für die entwicklungspolitische*. Freiburg i. Br.: Arnold Bergstraesser Institut.
- ODA. (1975). *A guide to the economic appraisal of projects in developing countries*. London: Stat. Off.
- OECD. (1986). *Methods and procedures in aid evaluation. A compendium of donor practice and experience*. Paris: Head of Publications Service, OECD.
- OECD. (1988). *Evaluation in developing countries. A step in a dialogue*. Paris: Head of Publications Service, OECD.
- OECD. (1991). *The DAC Principles for the Evaluation of Development Assistance*. Paris: OECD. Zugriff am 13. Juni 2019 unter <https://www.oecd.org/development/evaluation/2755284.pdf>
- OECD. (1993). *A History of the DAC Expert Group on Aid Evaluation* (OECD, Hrsg.) (Selected Issues in Aid Evaluation Nr. 7), Paris. Zugriff am 16. März 2019 unter <https://www.oecd.org/dac/evaluation/dcdndep/40163156.pdf>
- OECD. (1998). *Best Practice Guidelines for Evaluation* (PUMA Policy Brief Nr. 5). Zugriff am 20. Juni 2019 unter <http://portal.publicpolicy.utoronto.ca/en/StudyReports/OECD/BestPracticeGuidelinesforEvaluation/Documents/Best%20Practice%20Guidelines%20for%20Evaluation.pdf>
- OECD. (2005/2008). *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*. Zugriff am 02. September 2018 unter <http://www.oecd.org/dac/effectiveness/34428351.pdf>

- OECD. (2009). *Glossary of Key Terms in Evaluation and Results Based Management. Glossar entwicklungspolitischer Schlüsselbegriffe aus den Bereichen Evaluierung und ergebnisorientiertes Management* (OECD PUBLICATIONS, Hrsg.), Paris. Zugriff am 26. Oktober 2018 unter <https://www.oecd.org/dac/evaluation/dcdndep/43184177.pdf>
- OECD. (2010a). *Evaluating Development Cooperation. Summary of Key Norms and Standards* (Second Edition). Zugriff am 29. November 2019 unter <https://search.oecd.org/development/evaluation/dcdndep/41612905.pdf>
- OECD. (2010b). *Quality Standards for Development Evaluation* (DAC Guidelines and Reference Series): OECD PUBLICATIONS. Zugriff am 29. November 2019 unter <https://www.oecd.org/development/evaluation/qualitystandards.pdf>
- OECD. (2010c). *Evaluation in Development Agencies* (Better Aid). Paris: OECD Publishing. Zugriff am 27. November 2019 unter https://www.oecd-ilibrary.org/development/evaluation-in-development-agencies_9789264094857-en
- OECD. (2013). *The DAC Network on Development Evaluation - 30 years of strengthening learning in development*. Paris: DAC Network on Development Evaluation, OECD.
- OECD. (2016). *Evaluation Systems in Development Co-operation. 2016 Review*. Paris: OECD Publishing.
- OECD. (2018, 14. Dezember). *Germany Mid-term Review, 7 November 2018, Berlin*. Zugriff am 20. Januar 2020 unter <https://www.oecd.org/dac/peer-reviews/Germany-2018-Mid-term-review.pdf>
- OECD. (2019a). *The High-Level Fora on Aid Effectiveness: A history*. Zugriff am 05. Dezember 2019 unter <https://www.oecd.org/dac/effectiveness/thehighlevelforaonaideffectivenessahistory.htm>
- OECD (2019b). *Measuring Distance to the SDG Targets 2019. An Assessment of Where OECD Countries Stand*. Paris: OECD Publishing.
- OECD. (2019c). *Better Criteria for Better Evaluation. Revised Evaluation Criteria, Definitions and Principles for Use*. Zugriff am 13. Januar 2020 unter <http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>
- OECD/UNDP. (2019). *Making Development Co-operation More Effective. 2019 Progress Report*. Paris: OECD Publishing.
- Opaschowski, H. W. (2003). Pädagogik der Freizeit: Historische Entwicklung und zukünftige Entwicklungsperspektiven. In R. Popp & M. Schwab (Hrsg.), *Pädagogik der Freizeit* (Basiswissen Pädagogik, Pädagogische Arbeitsfelder, 6, S. 13-26). Baltmannsweiler: Schneider Verlag Hohengehren.
- Orth, M. (2019). Wie lässt sich Dauerhaftigkeit messen? In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 90). Bonn: DEval.
- Oxford, S. & Spaaij, R. (2017). Critical pedagogy and power relations in sport for development and peace: lessons from Colombia. *Third World Thematics: A TWQ Journal*, 2 (1), 102-116. doi: 10.1080/23802014.2017.1297687.

- Parthey, H. (2011). Institutionalisierung disziplinärer und interdisziplinärer Forschungssituationen. In K. Fischer, H. Latiko & H. Parthey (Hrsg.), *Interdisziplinarität und Institutionalisierung der Wissenschaft* (Wissenschaftsforschung, 2010, S. 9-36). Berlin: Wissenschaftlicher Verlag.
- Partnerschaften 2030 (Deutsche GIZ GmbH, Hrsg.). (o.D.). *Was ist eine MAP*. Zugriff am 23. März 2021 unter <https://www.partnerschaften2030.de/was-ist-eine-map/>
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, Calif.: Sage.
- Patton, M. Q. (2011). *Developmental evaluation. Applying complexity concepts to enhance innovation and use*. New York, NY: Guilford Press.
- Patton, M. Q. (2019). *Blue marble evaluation. Premises and principles*. New York: The Guilford Press.
- Pawson, R. & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- Petry, K. (Hrsg.). (2020). *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage). Leverkusen: Verlag Barbara Budrich.
- Petry, K. (2020). *Sport und Entwicklung: Eine aktualisierte Übersicht über ausgewählte in Deutschland ansässige Akteure und ihrer Aktivitäten. Zusammengestellt im Auftrag der Deutschen Gesellschaft für Internationale Zusammenarbeit GmbH*. Interner Bericht: Deutsche Sporthochschule Köln.
- Petry, K. & Bauer, K. (2016). *Schlussbericht Wissenschaftliche Begleitung des Sektorvorhabens "Sport für Entwicklung" der GIZ. Berichtszeitraum: 15.07.2013-30.06.2016*. Interner Bericht: Deutsche Sporthochschule Köln.
- Petry, K. & Bauer, K. (2019). *Schlussbericht Wissenschaftliche Begleitung des Sektorvorhabens "Sport für Entwicklung" der GIZ. Berichtszeitraum: 1.07.2016-30.09.2019*. Interner Bericht: Deutsche Sporthochschule Köln.
- Petry, K., Bauer, K. & Biermann, M. (2020). Wirkungsmessung bei sozialen Sportprojekten im Kontext der Entwicklungszusammenarbeit: Methodische Überlegungen und Herausforderungen. In K. Petry (Hrsg.), *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage, S. 221-230). Leverkusen: Verlag Barbara Budrich.
- Petry, K., Bauer, K., Dumon, D. & Wyszynski, M. (2015). *Dokumentation und Analyse von ausgewählten Programmen der deutschen Entwicklungszusammenarbeit im Bereich „Sport für Entwicklung“*. Interner Bericht: DSHS Köln; ICSSPE Berlin.
- Petry, K., Groll, M. & Tokarski, W. (Hrsg.). (2011). *Sport und internationale Entwicklungszusammenarbeit. Theorie- und Praxisfelder* (Veröffentlichungen der Deutschen Sporthochschule Köln, Bd. 17, 1. Aufl.). Köln: Sportverlag Strauß.
- Petry, K. & Neeten, M. (Dezember 2019). *Evaluation report Tool 4 Coaches and Instructors*. Interner Bericht: DSHS Köln.
- Petry, K., Pedemonte, L., Schröder, R. & Bauer, K. (2014). *Akteursanalyse: Friedenförderung durch Sport in Kolumbien*. Interner Bericht: DSHS Köln.

- Popp, R. & Schwab, M. (2003). Von der Freizeitpädagogik zur Pädagogik der Freizeit. In R. Popp & M. Schwab (Hrsg.), *Pädagogik der Freizeit* (Basiswissen Pädagogik, Pädagogische Arbeitsfelder, 6, S. 1-12). Baltmannsweiler: Schneider Verlag Hohengehren.
- Prenzlau, J. & Rordorf, I. (2008). *Die Entwicklung qualitativer Instrumente zur Erfassung der Wirksamkeit einer Interventionsmaßnahme zur interkulturellen Verständigung in Israel*. Diplomarbeit, Deutsche Sporthochschule Köln.
- Preti, D. (2012). Monitoring and Evaluation. Between the Claims and Reality. In K. Gilbert & W. Bennett (Eds.), *Sport, peace and development* (pp. 309-317). Champaign, Illinois: Common Ground Pub. LLC.
- Priest, S. & Gass, M. A. (2005). *Effective leadership in adventure programming* (2nd ed.). Champaign, Ill.: Human Kinetics.
- Rabie, B. (2014). Evaluation Models, Theories and Paradigms. In F. Cloete, B. Rabie & C. de Coning (Hrsg.), *Evaluation management in South Africa and Africa* (First edition, S. 116-151). Stellenbosch: SUN Media.
- Rabie, B. & Goldman, I. (2014). The Context of Evaluation Management. In F. Cloete, B. Rabie & C. de Coning (Hrsg.), *Evaluation management in South Africa and Africa* (First edition, S. 3-27). Stellenbosch: SUN Media.
- Ramboll Management Consulting. (2018). *What's the bussiness potential of the UN SDGs?* Zugriff am 22. März 2021 unter <http://consulting.ramboll.com/action/media/18558/sdg-assessment-tool>
- Reade, N. (2008). *Konzept für alltagstaugliche Wirkungsevaluierungen in Anlehnung an Rigorous Impact Evaluations* (CEval-Arbeitspapiere Nr. 14). Saarbrücken: Centrum für Evaluation.
- Reiners, A. (2007). *Bewährte Sammlung motivierender Interaktionsspiele* (Gelbe Reihe: Praktische Erlebnispädagogik, 8., überarbeitete Auflage). Augsburg: ZIEL.
- Reiners, A. (2014). *Neue Sammlung handlungsorientierter Übungen für Seminar und Training* (Gelbe Reihe, Band 2, 2. überarbeitete Auflage 2007, Nachdruck 2014). Augsburg: ZIEL.
- Reis, A. C., Vieira, M. C. & Sousa-Mast, F. R. de. (2016). "Sport for Development" in developing countries: The case of the Vilas Olímpicas do Rio de Janeiro. *Sport Management Review*, 19 (2), 107-119. doi: 10.1016/j.smr.2015.01.005.
- Reuber, M. (2006). Evaluierung in der Gesellschaft für Technische Zusammenarbeit (GTZ). *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 67-94.
- Richelieu, A. & Webb, A. (2019). A Tale of Two Networks: Contrasting Sport for Development and Peace Actor-Networks. *Journal of Global Sport Management*, 4 (4), 371-391. doi: 10.1080/24704067.2018.1477517.
- Riddell, R. C. (1987). *Foreign aid reconsidered*. Baltimore: Johns Hopkins Univ. Pr.
- Riddell, R. C. (2002). *Synthesis Report on Development Agency Policies and Perspectives on Programme-Based Approaches. Prepared for the Forum on Accountability and Risk Management Under Programme-Based Approaches*, Ottawa.

- Rittner, V. (1977). Sport als Kulturexport. *Kulturaustausch: Zeitschrift für internationale Perspektiven*, 27 (4), 34-41.
- Rockenfeller, S. (2009). *Namibia Sport Projekt 2008 - Möglichkeiten und Grenzen für die Entwicklungszusammenarbeit im Sport. Eine qualitative Analyse*. Diplomarbeit, Deutsche Sporthochschule Köln. Köln.
- Rodrick, D. & Rosenzweig, M. R. (2009). *Handbook of development economics. Volume 5*. Burlington: Elsevier.
- Rodríguez-Bilella, P. & Lucero, M. A. (2017). Evaluation als globales Phänomen: Die Entwicklung transnationaler Netzwerke. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 53-70). Münster: Waxmann Verlag GmbH.
- Rolff, K. (2020). *Sport und Kompetenzentwicklung. Mannschaftssport als Instrument zur Steigerung von Kompetenzen junger bildungsferner Menschen*. Zugriff am 18. November 2020 unter <https://www.iat.eu/aktuell/veroeff/jahrbuch/jahrb07/13-rolff.pdf>
- Rossi, P. H., Lipsey, M. W. & Freeman, H. E. (2004). *Evaluation. A systematic approach* (7. ed.). Thousand Oaks, California: Sage.
- Roth, K. & Kröger, C. (2015). *Ballschule. Ein ABC für Spielanfänger* (Praxisideen Sportspiele, Band 1, 5. Auflage). Schorndorf: Hofmann.
- Röthig, P. & Prohl, R. (Hrsg.). (2003). *Sportwissenschaftliches Lexikon* (Beiträge zur Lehre und Forschung im Sport, 49/50, 7., völlig neu bearb. Aufl.). Schorndorf: Hofmann.
- Roxin, H. (2016). Die Wirksamkeit der Entwicklungszusammenarbeit auf dem Prüfstand. Eine kritische Analyse: Evaluation als Kernaufgabe zukünftiger Entwicklungszusammenarbeit. *Außerschulische Bildung*, 2, 12-18.
- Roxin, H. & Asche, H. (2013). DEval. *Zeitschrift für Evaluation (ZfEv)*, 12 (2), 297-304.
- RTP. (2008). *Harnessing the Power of Sport for Development and Peace: Recommendations to Governments: Sport for Development and Peace International Working Group (SDP IWG)*. Zugriff am 20. Mai 2020 unter https://www.sportand-dev.org/sites/default/files/downloads/45b__harnessing_the_power_of_sport_for_development_and_peace__report_from_the_athens_.pdf
- Rudolph, A. & Holzapfel, S. (2017). *Drei Schritte die Wirksamkeit von EZ mit der Agenda 2030 zu verknüpfen* (Dt. Institut für Entwicklungspolitik (DIE), Hrsg.) (Die aktuelle Kolumne vom 06.03.2017).
- S&P Trucost Limited. (2021). *The Trucost SDG Evaluation Tool*. Zugriff am 22. März 2021 unter: <https://www.trucost.com/corporate-advisory/sdg-evaluation-tool/>
- S4D Resource Toolkit. (o.D.). *Home*. Zugriff am 08. Januar 2022 unter <https://www.sport-for-development.com/home>
- S4D Resource Toolkit. (o.D.a). *Country Collection*. Zugriff am 18. Oktober 2020 unter <https://www.sport-for-development.com/country-collection>
- S4D Resource Toolkit. (o.D.b). *S4D Partners and Networks*. Zugriff am 02. November 2020 unter <https://www.sport-for-development.com/essentials?id=18#cat18>

- S4D Resource Toolkit. (o.D.c). *S4D Tools. S4D Teaching and Learning Materials*. Zugriff am 18. November 2020 unter <https://www.sport-for-development.com/tools?id=34#cat34>
- S4D Resource Toolkit. (o.D.d). *S4D Competences, Life Skills* unter <https://www.sport-for-development.com/essentials?id=2#cat2>
- Saavedra, M. (2019). SDP and postcolonial theory. In H. Collison, S. C. Darnell, R. Giuliannotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 208-217). Milton: Routledge.
- Salmen, L. F. (1987). *Listen to the people. Participant-observer evaluation of development projects* (1. print). New York u.a.: Oxford University Press.
- Sanders, B. (2019, 23. April). *Closing the Gap - Measuring the contribution of sport to the SDGs*. Zugriff am 20. Mai 2020 unter <https://www.sportanddev.org/en/article/news/closing-gap-measuring-contribution-sport-sdgs>
- Sanders, B. (2020, 25. August). *Online course attracts students from 153 countries*. Zugriff am 29. März 2021 unter <https://www.sportanddev.org/en/article/news/online-course-attracts-students-153-countries>
- Sangmeister, H. & Wagner, H. (Hrsg.). (2019). *Die Entwicklungszusammenarbeit der Zukunft*. Baden-Baden: Nomos.
- Sasaki, R. (2006). A Review of the History and the Current Practice of Aid Evaluation. *Journal of Multi Disciplinary Evaluation*, 3 (5), 55-88. Zugriff am 15. März 2019 unter http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/51
- Savedoff, W. D., Levine, R. & Birdsall, N. (2006). *When Will We Ever Learn? Improving Lives through Impact Evaluation* (Center for Global Development (CGD), Hrsg.). Zugriff am 30. November 2019 unter https://www.cgdev.org/sites/default/files/7973_file_WillWeEverLearn.pdf
- Schemme, D. (2003). Wissenschaftliche Begleitung und Evaluation in Wirtschaftsmodellversuchen. *Berufsbildung in Wissenschaft und Praxis* (6), 31-35.
- Schmid, K. (2015). *Weiterbildungsmanagement in der sportbezogenen Entwicklungszusammenarbeit. Analyse einer Trainerausbildungsmaßnahme der GIZ in Brasilien*. Bachelorarbeit, Deutsche Sporthochschule Köln. Köln.
- Schmid, K. & Arnemann, A. (2020). Sport für Entwicklung in Afrika - Ansätze und Erfahrungen der deutschen Entwicklungszusammenarbeit und ihrer Partner. In K. Petry (Hrsg.), *Sport im Kontext von internationaler Zusammenarbeit und Entwicklung. Perspektiven und Herausforderungen im Spannungsfeld von Wissenschaft, Politik und Praxis* (1. Auflage, S. 155-166). Leverkusen: Verlag Barbara Budrich.
- Schmidt, H. R. & Florin, H. (1989). Überblick der Ausbildung und Hilfsmaßnahmen des Deutschen Fußball-Bundes e.V. In R. Andresen, H. Rieder & G. Trosien (Hrsg.), *Beiträge zur Zusammenarbeit im Sport mit der Dritten Welt* (Texte, Quellen, Dokumente zur Sportwissenschaft, 23, S. 170-176). Schorndorf: Hofmann.
- Schmitt, J. (2018). *Schluss mit schwarzen Boxen. Zur Arbeit mit Kausalmechanismen in Evaluierungen*. DEval Policy Brief 10/2018: DEval.

- Schneider, S. & Eger, J. (2019). Die Agenda 2030 aus Sicht der Bevölkerung. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 64-65). Bonn: DEval.
- Schneider, S. H., Eger, J. & Sassenhagen, N. (2021). *Meinungsmonitor Entwicklungspolitik 2021. Medieninhalte, Informationen, Appelle und ihre Wirkung auf die öffentliche Meinung*. Bonn: DEval.
- Schneider, S. H. & Gleser, S. H. (2018). *Meinungsmonitor Entwicklungspolitik 2018: Einstellungen zu Entwicklungszusammenarbeit und nachhaltiger Entwicklung*. Bonn: DEval.
- Schnell, R., Hill, P. B. & Esser, E. (1999). *Methoden der empirischen Sozialforschung* (9., aktualisierte Aufl.). München: Oldenbourg.
- Schödlbauer, C. (2000). *Metaphorisches Lernen in erlebnispädagogischen Szenarien. Eine Untersuchung über handlungsorientierte Lehr-Lern-Prozesse* (Schriftenreihe EUB, Erziehung - Unterricht - Bildung, 86). Hamburg: Kovač.
- Scholz, I. & Wolff, P. (2010). Was bedeutet Wirksamkeit für die Technische und Finanzielle Entwicklungszusammenarbeit? In J. Faust & S. Neubert (Hrsg.), *Wirksamere Entwicklungspolitik. Befunde, Reformen, Instrumente* (1. Auflage, S. 320-345). Baden-Baden: Nomos.
- Schott, N. & Merkel, K. (2009). Chancen und Wege der Entwicklungszusammenarbeit im Sport. In E. Balz & D. Kuhlmann (Hrsg.), *Sportentwicklung. Grundlagen und Facetten* (Sportentwicklungen in Deutschland, 21, S. 119-140). Aachen: Meyer & Meyer.
- Schreiber, J.-R. & Siege, H. (2016). *Orientierungsrahmen für den Lernbereich Globale Entwicklung. im Rahmen einer Bildung für nachhaltige Entwicklung* (2. aktualisierte und erweiterte Auflage) (Kultusministerkonferenz (KMK) & Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Hrsg.). Zugriff am 10. Dezember 2020 unter https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_00-Orientierungsrahmen-Globale-Entwicklung.pdf
- Schreiner, L., Mayer, J. & Kastrup, V. (2020). National 'Sport for Development and Peace'-programmes in light of UN recommendations and national sports organisations' endeavours. A document analysis of the role of 'Sports' in Germany's development policy. *European Journal for Sport and Society*, 1-21. doi: 10.1080/16138171.2020.1823691.
- Schröter, D. (2020). Evaluation of Sustainability for Sustainability: The Sustainability Evaluation Checklist Revisited. *Zeitschrift für Evaluation (ZfEv)*, 19 (2), 291-320.
- Schulenkorf, N. (2012). Sustainable community development through sport and events: A conceptual framework for Sport-for-Development projects. *Sport Management Review*, 15 (1), 1-12. doi: 10.1016/j.smr.2011.06.001.
- Schulenkorf, N. (2013). Sport-for-Development Events and Social Capital Building: A Critical Analysis of Experiences from Sri Lanka. *Journal of Sport for Development*, 1 (1), 25-36.

- Schulenkorf, N. (2017). Managing sport-for-development: Reflections and outlook. *Sport Management Review*, 20 (3), 243-251. doi: 10.1016/j.smr.2016.11.003.
- Schulenkorf, N., Edwards, D. & Hergesell, A. (2020). Guiding qualitative inquiry in sport-for-development: The sport in development settings (SPIDS) research framework. *Journal of Sport for Development*, 8 (14), 53-69.
- Schulenkorf, N., Sherry, E. & Rowe, K. (2016). Sport for Development: An Integrated Literature Review. *Journal of Sport Management*, 30 (1), 22-39. doi: 10.1123/jsm.2014-0263.
- Schulenkorf, N. & Siefken, K. (2019). Managing sport-for-development and healthy lifestyles: The sport-for-health model. *Sport Management Review*, 22 (1), 96-107. doi: 10.1016/j.smr.2018.09.003.
- Schulenkorf, N. & Sugden, J. (2011). Sport for Development and Peace in Divided Societies: Cooperating for Inter-Community Empowerment in Israel. *European Journal for Sport and Society*, 8 (4), 235-256. doi: 10.1080/16138171.2011.11687881.
- Schulze, S. (2021). *Über mich*. Zugriff am 03. Februar 2022 unter <https://svenja-schulze.de/ueber-mich/>
- Schweitzer. (2013). *Policy für Monitoring und Evaluierung der GIZ*. Bonn, Eschborn: GIZ GmbH. Zugriff am 13. März 2020 unter https://www.giz.de/de/downloads/giz2013-de-Policy_Monitoring_Evaluierung.pdf
- Scriven, M. (2008). *Evaluation thesaurus* (4. Aufl.). Newbury Park, Calif: Sage Publications.
- SDG Fund. (2018). *The Contribution of Sports to the Achievement of Sustainable Development Goals: A Toolkit for Action*. Monaco: UN Department of Economic and Social Affairs (UN DESA). Zugriff am 20. Mai 2020 unter <https://digitallibrary.un.org/record/1657855/files/Sdgsports.pdf>
- SDP IWG. (2007). *Literature Reviews on Sport for Development and Peace* (Sport for Development and Peace International Working Group (SDP IWG) Secretariat, Hrsg.): University of Toronto, Canada.
- SDP IWG. (2008). *Harnessing the power of sport for development and peace: Recommendations to governments*. Toronto, Canada: Right to Play. Zugriff am 20. Januar 2020 unter https://www.sportanddev.org/sites/default/files/downloads/rtp_sdp_iwg_harnessing_the_power_of_sport_for_development_and_peace.pdf
- Segone, M., Heider, C., Oksanen, R., Silva, S. de & Sanz, B. (2013). Towards a shared framework for National Evaluation Capacity Development. In M. Segone & J. Rugh (Hrsg.), *Evaluation and Civil Society Stakeholders' perspectives on National Evaluation Capacity Development* (S. 16-42). New York: UNICEF.
- SEVAL. (2016). *Evaluationsstandards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards)* (SEVAL, Hrsg.). Zugriff am 19. Juni 2021 unter https://www.seval.ch/app/uploads/2018/01/SEVAL-Standards-2016_d.pdf
- Shadish, W. R., Cook, T. D. & Leviton, L. C. (Eds.). (1991). *Foundations of program evaluation. Theories of practice* (Reprinted.). Newbury Park: Sage Publications.

- Shaw, I., Greene, J. C. & Mark, M. M. (Eds.). (2007). *Handbook of evaluation. Policies, programs and practices*. London: Sage Publ.
- Shediac-Rizkallah, M. C. & Bone, L. R. (1998). Planning for the sustainability of community-based health programs: conceptual frameworks and future directions for research, practice and policy. *Health Education Research*, 13 (1), 87-108.
- Sherry, E. & Osborne, A. (2019). SDP and homelessness. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 374-384). Milton: Routledge.
- Sherry, E., Schulenkorf, N. & Phillips, P. (2016). Evaluating sport development. In E. Sherry, N. Schulenkorf & P. Phillips (Hrsg.), *Managing sport development. An international approach* (S. 161-176). New York: Routledge.
- Sherry, E., Schulenkorf, N., Seal, E., Nicholson, M. & Hoye, R. (2017a). Sport-for-Development in the South Pacific Region: Macro-, Meso-, and Micro-Perspectives. *Sociology of Sport Journal*, 34 (4), 303-316. doi: 10.1123/ssj.2017-0022.
- Sherry, E., Schulenkorf, N., Seal, E., Nicholson, M. & Hoye, R. (2017b). Sport-for-development: Inclusive, reflexive, and meaningful research in low- and middle-income settings. *Sport Management Review*, 20 (1), 69-80. doi: 10.1016/j.smr.2016.10.010.
- Shin, N., Cohen, A. & Welty Peachey, J. (2020). Advancing the sport for development field: Perspectives of practitioners on effective organizational management. *Journal of Sport for Development*, 8 (14), 36-52.
- Silvestrini, S. & Reade, N. (2008). *CEval-Ansatz zur Wirkungsevaluation / Stockmann'scher Ansatz* (Centrum für Evaluation (CEval), Hrsg.) (CEval-Arbeitspapiere Nr. 11), Saarbrücken.
- Skinner, J., Woolcock, G. & Yoon, L. (2019). SDP and social capital. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 296-307). Milton: Routledge.
- Skinner, J., Zakus, D. H. & Cowell, J. (2008). Development through Sport: Building Social Capital in Disadvantaged Communities. *Sport Management Review*, 11 (3), 253-275. doi: 10.1016/S1441-3523(08)70112-8.
- Sonnenschein, W. & Müller, H.-J. (1996). *Sport und Sportförderung in Entwicklungsländern. Eine Bibliographie* (Berichte und Materialien des Bundesinstituts für Sportwissenschaft, 96,11, 1. Aufl.). Köln: Sport und Buch Strauss.
- Spaaij, R. (2009). The social impact of sport: diversities, complexities and contexts. *Sport in Society*, 12 (9), 1109-1117. doi: 10.1080/17430430903137746.
- Spaaij, R. & Jeanes, R. (2013). Education for social change? A Freirean critique of sport for development and peace. *Physical Education & Sport Pedagogy*, 18 (4), 442-457. doi: 10.1080/17408989.2012.690378.
- Spaaij, R., Schulenkorf, N., Jeanes, R. & Oxford, S. (2018). Participatory research in sport-for-development: Complexities, experiences and (missed) opportunities. *Sport Management Review*, 21 (1), 25-37. doi: 10.1016/j.smr.2017.05.003.

- Spiel, C. & Schober, B. (2015). Der Beitrag wissenschaftlicher Evaluation zur Evidenzbasiierung in Politik und Praxis. In V. Hennefeld, W. Meyer, S. Silvestrini & R. Stockmann (Hrsg.), *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft: Festschrift zum 60. Geburtstag von Reinhard Stockmann* (S. 41-52). Münster: Waxmann.
- Sportanddev.org. (o.D.). *Timeline of major developments in sport and development*. Zugriff am 23. Mai 2020 unter <https://www.sportanddev.org/fr/node/112>
- Steinbach, D., Groll, M., Haasner, O. & Damerius, I. (2008). *Zukunft - Freizeit - Sport: Situation und Perspektiven des Freizeit- und Breitensports in Salzburg*. Salzburg: Fachhochschule Salzburg, Forschungsges. mbH.
- Steinbach, D. & O'Rourke, B. (2017). *Learn to Play - Play to Learn. A Practical Handbook for Life Skills Coaches in Sport for Development*. Bonn, Eschborn: GIZ GmbH.
- Steinmann, W. (2011). Auslandstrainerschule Mainz - ein Beitrag zur internationalen Sportförderung des Auswärtigen Amtes. In K. Petry, M. Groll & W. Tokarski (Hrsg.), *Sport und internationale Entwicklungszusammenarbeit. Theorie- und Praxisfelder* (S. 135-146). Köln: Sportverlag Strauß.
- Stockmann, R. (1992). *Die Nachhaltigkeit von Entwicklungsprojekten. Eine Methode zur Evaluierung am Beispiel von Berufsbildungsprojekten*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stockmann, R. (1996). *Die Wirksamkeit der Entwicklungshilfe. Eine Evaluation der Nachhaltigkeit von Programmen und Projekten der Berufsbildung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stockmann, R. (2000a). Evaluation in Deutschland. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 11-40). Opladen: Leske + Budrich.
- Stockmann, R. (2000b). Evaluation staatlicher Entwicklungspolitik. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 375-407). Opladen: Leske + Budrich.
- Stockmann, R. (2006). *Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement* (Sozialwissenschaftliche Evaluationsforschung, 5). Münster: Waxmann.
- Stockmann, R. (2007a). Einleitung: Wozu brauchen wir Evaluation? In R. Stockmann (Hrsg.), *Handbuch zur Evaluation. Eine praktische Handlungsanleitung* (Sozialwissenschaftliche Evaluationsforschung, 6, S. 13-23). Münster: Waxmann.
- Stockmann, R. (2007b). Einführung in die Evaluation. In R. Stockmann (Hrsg.), *Handbuch zur Evaluation. Eine praktische Handlungsanleitung* (Sozialwissenschaftliche Evaluationsforschung, 6, S. 24-70). Münster: Waxmann.
- Stockmann, R. (2007c). Konkurrierende und komplementäre Ansätze zur Evaluation C. In R. Stockmann (Hrsg.), *Handbuch zur Evaluation. Eine praktische Handlungsanleitung* (Sozialwissenschaftliche Evaluationsforschung, 6, S. 71-107). Münster: Waxmann.
- Stockmann, R. (2012). Von der Idee zur Institution. Institut für deutsche Entwicklungsevaluierung gegründet. *Zeitschrift für Evaluation (ZfEv)*, 11 (1), 85-93.

- Stockmann, R. (2017). Die Zukunft der Evaluation in modernen Gesellschaften. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 35-52). Münster: Waxmann Verlag GmbH.
- Stockmann, R. (2018). Der Evaluation Capacity Development Ansatz des Centrums für Evaluation (CEval). *Zeitschrift für Evaluation (ZfEv)*, 17 (2), 363-392.
- Stockmann, R. & Borrmann, A. (2009). Das Evaluationssystem in der deutschen Entwicklungszusammenarbeit. *Zeitschrift für Evaluation (ZfEv)*, 8 (2), 263-295.
- Stockmann, R. & Gaebe, W. (Hrsg.). (1993). *Hilft die Entwicklungshilfe langfristig? Bestandsaufnahme zur Nachhaltigkeit von Entwicklungsprojekten*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stockmann, R., Menzel, U. & Nuscheler, F. (2016). *Entwicklungspolitik. Theorien - Probleme - Strategien* (2., aktualisierte und überarbeitete Auflage). Berlin: De Gruyter Oldenbourg.
- Stockmann, R. & Meyer, W. (2017a). Evaluation in Deutschland. Woher sie kommt, wo sie steht, wohin sie geht. *Zeitschrift für Evaluation (ZfEv)*, 16 (2), 57-110.
- Stockmann, R. & Meyer, W. (2017a). Einleitung. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 9-20). Münster: Waxmann Verlag GmbH.
- Stockmann, R. & Meyer, W. (2017b). Evaluation zwischen Wissenschaft und Praxis. In R. Stockmann & W. Meyer (Hrsg.), *Die Zukunft der Evaluation. Trends, Herausforderungen, Perspektiven* (1. Aufl., S. 155-176). Münster: Waxmann Verlag GmbH.
- Stockmann, R. & Meyer, W. (2020). Die Institutionalisierung der Evaluation in Europa. *Zeitschrift für Evaluation (ZfEv)*, 19 (1), 111-135.
- Stockmann, R., Meyer, W. & Schenke, H. (2011). Unabhängigkeit von Evaluationen. *Zeitschrift für Evaluation (ZfEv)*, 10 (1), 39-67.
- Stokke, O. (1991). Policies, Performance, Trends and Challenges in Aid Evaluation. In O. Stokke (Ed.), *Evaluating development assistance. Policies and performance* (EADI book series, vol. 12, pp. 1-59). London: Cass.
- Strange, T. & Bayley, A. (2008). *Nachhaltige Entwicklung. Wirtschaft, Gesellschaft, Umwelt in Zusammenhang betrachtet* (OECD Insights). Paris: OECD PUBLICATIONS.
- Sugden, J. T., Adair, D., Schulenkorf, N. & Frawley, S. (2019). Exploring Sport and Inter-group Relations in Fiji: Guidance for Researchers Undertaking Short-Term Ethnography. *Sociology of Sport Journal*, 36 (4), 277-288 unter <https://researchonline.ljmu.ac.uk/id/eprint/13641/>
- Sugden, J. (2010). Critical left-realism and sport interventions in divided societies. *International Review for the Sociology of Sport*, 45 (3), 258-272. doi: 10.1177/1012690210374525.
- Sugden, J. (2014). The Ripple Effect: Critical Pragmatism, Conflict Resolution and Peace Building through Sport in Deeply Divided Societies. In N. Schulenkorf & D. Adair (Eds.), *Global Sport-for-Development. Critical Perspectives* (pp. 79-98). Basingstoke: Palgrave Macmillan.

- Svensson, P. G. & Hambrick, M. E. (2019). Exploring how external stakeholders shape social innovation in sport for development and peace. *Sport Management Review*, 22 (4), 540-552. doi: 10.1016/j.smr.2018.07.002.
- Svensson, P. G., Hancock, M. G. & Hums, M. A. (2016). Examining the educative aims and practices of decision-makers in sport for development and peace organizations. *Sport, Education and Society*, 21 (4), 495-512. doi: 10.1080/13573322.2015.1102723.
- Svensson, P. G. & Levine, J. (2017). Rethinking Sport for Development and Peace: The Capability Approach. *Sport in Society*, 20 (7), 905-923. doi: 10.1080/17430437.2016.1269083.
- Svensson, P. G. & Woods, H. (2017). A systematic overview of sport for development and peace organizations. *Journal of Sport for Development*, 5 (9), 36-48.
- Swiss, J. E. (2005). A Framework for Assessing Incentives in Results-Based Management. *Public Administration Review*, 65 (5), 592-602. doi: 10.1111/j.1540-6210.2005.00486.x.
- Terberger, E. (2011). Evaluierung in der Entwicklungszusammenarbeit. Das Beispiel der Finanziellen Zusammenarbeit. In J. König & J. Thema (Hrsg.), *Nachhaltigkeit in der Entwicklungszusammenarbeit. Theoretische Konzepte, strukturelle Herausforderungen und praktische Umsetzung* (Globale Gesellschaft und internationale Beziehungen, 1. Aufl., S. 219-238). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Thorpe, H. (2016). 'Look at what we can do with all the broken stuff!' Youth agency and sporting creativity in sites of war, conflict and disaster. *Qualitative Research in Sport, Exercise and Health*, 8 (5), 554-570. doi: 10.1080/2159676X.2016.1206957.
- Tiessen, R. (2011). Global Subjects or Objects of Globalisation? The promotion of global citizenship in organisations offering sport for development and/or peace programmes. *Third World Quarterly*, 32 (3), 571-587. doi: 10.1080/01436597.2011.573946.
- Tokarski, W. & Petry, K. (Hrsg.). (2010). *Handbuch Sportpolitik* (Beiträge zur Lehre und Forschung im Sport, Bd. 172). Schorndorf: Hofmann.
- Treichel, K., Höh, A., Biermann, S. & Conze, P. (2016). *Multi-Akteurs-Partnerschaften im Rahmen der Agenda 2030. Eine praxisorientierte Analyse von Potentialen, Herausforderungen und Erfolgsfaktoren* (Partnerschaften 2030 c/o GIZ GmbH, Hrsg.), Bonn. Zugriff am 12. März 2021 unter https://www.partnerschaften2030.de/wp-content/uploads/2017/09/Viadrina_Governance_MAP-Studie_WEB.pdf
- Trosien, G. (1987). *Sportorganisationen und Dritte Welt. Kompetenzen, Konzepte, korporative Strukturen* (1. Aufl.). Sankt Augustin: Richarz.
- Tuohey, B. & Cognato, B. (2011). PeacePlayers International: A case study on the use of sport as a tool for conflict transformation. *SAIS Review of International Affairs*, 31, 31 (1), 51-63.

- UN DESA. (2019). *2019 Voluntary National Reviews. Synthesis Report*, High-level Political Forum on Sustainable Development. Zugriff am 15. Januar 2020 unter https://sustainabledevelopment.un.org/content/documents/252302019_VNR_Synthesis_Report_DESA.pdf
- UN DESA. (2019, November). *Handbook for the Preparation of Voluntary National Reviews. The 2020 Edition*, High-level Political Forum on Sustainable Development. Zugriff am 15. Januar 2020 unter https://sustainabledevelopment.un.org/content/documents/25245Handbook_2020_EN.pdf
- UN General Assembly. (1959). *Declaration of the Rights of the Child. Resolution A/RES/14/1386*. Zugriff am 18. Oktober 2020 unter [https://www.ohchr.org/EN/Issues/Education/Training/Compilation/Pages/1DeclarationoftheRightsoftheChild\(1959\).aspx](https://www.ohchr.org/EN/Issues/Education/Training/Compilation/Pages/1DeclarationoftheRightsoftheChild(1959).aspx)
- UN General Assembly. (2003). *Sport as a means to promote education, health, development and peace. Resolution A/RES/58/5*, New York. Zugriff am 20. Januar 2020 unter https://www.sportanddev.org/sites/default/files/downloads/35__un_general_assembly_resolution_58_5__sport_as_a_means_to_promote_health_educatio.pdf
- UN General Assembly. (2015a). *Transforming our world: the 2030 Agenda for Sustainable Development. Resolution A/RES/70/1*. Zugriff am 31. Januar 2020 unter https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/70/1
- UN General Assembly. (2015b). Draft outcome document of the United Nations summit for the adoption of the post-2015 development agenda. A/69/L.85 unter https://www.un.org/ga/search/view_doc.asp?symbol=A/69/L.85&Lang=E
- UN General Assembly. (2018). *Sport as an enabler of sustainable development. Resolution A/RES/73/24*. Zugriff am 18. Mai 2020 unter https://digitallibrary.un.org/record/1655098/files/A_RES_73_24-EN.pdf
- UNDP. (2022). *2019 Global Progress Report*, United Nations Development Programme. Zugriff am 19. Februar 2022 unter <https://www.effectivecooperation.org/landing-page/2018-monitoring-results>
- UNEG. (2011). *Integrating Human Rights and Gender Equality in Evaluation. Towards UNEG Guidance*. Zugriff am 12. März 2021 unter <http://www.uneval.org/document/download/1294>
- UNEG. (2013, 14. Mai). *Subject: UNEG joins EvalPartners in declaring 2015 as International Year of Evaluation*. Zugriff am 20. Januar 2020 unter <http://www.unevaluation.org/document/detail/1378>
- UNEG. (2015, 25. September). *Evaluation fit for the 2030 Sustainable Development Agenda: A commitment by evaluators*. Zugriff am 20. Januar 2020 unter <http://www.uneval.org/mediacenter/newscenter/newsdetail/114>
- UNESCO. (1978). *UNESCO International Charter of Physical Education and Sport*. Zugriff am 18. Oktober 2020 unter https://www.sportanddev.org/sites/default/files/downloads/17__intl_charter_of_pe_and_sport.pdf

- UNESCO. (2017). *Final Report: Sixth International Conference of Ministers and Senior Officials responsible for Physical Education and Sport (MINEPS VI), 14-15 July 2017*
- UNESCO (2021). *Fit for Life: sport powering inclusive, peaceful, and resilient societies*. Zugriff am 30. März 2022 unter <https://unesdoc.unesco.org/ark:/48223/pf0000379910>
- Kazan, Russia: United Nations Educational Scientific and Cultural Organization (UNESCO).
- United Nations Inter-Agency Task Force on Sport for Development and Peace. (2003). *Sport for Development and Peace. Towards Achieving the Millennium Development Goals*. Zugriff am 07. Oktober 2020 unter <https://www.unicef.org/sports/reportE.pdf>
- UNOSDP. (2006). *Report on the International Year of Sport and Physical Education 2005. Sport for a Better World*. Geneva: Publishing Service United Nations.
- Vaessen, J. (2019). Nutzung einer systemischen Perspektive in Evaluierungen. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 30-31). Bonn: DEval.
- Valadez, J. & Bamberger, M. (1994). *Monitoring and evaluating social programs in developing countries: a handbook for policymakers, managers, and researchers*. Washington D.C.: World Bank.
- Van Eekeren, F., ter Horst, K. & Fictorie, D. (2013). *Sport for Development: The potential value and next steps. Reviw of policy, programs and academic research 1998-2013*. Arnhem: Foundation LM Publishers.
- Vedung, E. (2000). Evaluation Research and Fundamental Research. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 103-126). Opladen: Leske + Budrich.
- Vedung, E. (2015). Six Uses of Evaluation. In V. Hennefeld, W. Meyer, S. Silvestrini & R. Stockmann (Hrsg.), *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft: Festschrift zum 60. Geburtstag von Reinhard Stockmann* (S. 187-210). Münster: Waxmann.
- Verspohl, I. & Hartmann, C. (2019). Länderportfolios: Mehr strategische Planung und Steuerung notwendig. In DEval (Hrsg.), *Nachhaltigkeit gestalten. Die Agenda 2030 in der Entwicklungszusammenarbeit* (S. 22-23). Bonn: DEval.
- Vestman, O. K. & Conner, R. F. (2007). The Relationship between Evaluation and Politics. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *Handbook of evaluation. Policies, programs and practices* (pp. 225-242). London: Sage Publ.
- Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation Practice*, 14 (1), 93-106. doi: 10.1016/0886-1633(93)90046-R.
- Weiss, C. H. (1997). How Can Theory-Based Evaluation Make Greater Headway? *Evaluation Review*, 21 (4), 501-524. doi: 10.1177/0193841X9702100405.
- Weiss, C. H. (1998). Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19 (1), 21-33. doi: 10.1177/109821409801900103.

- Welty Peachey, J. (2019). SDP and sport management. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 241-252). Milton: Routledge.
- Welty Peachey, J., Bruening, J., Lyras, A., Cohen, A. & Cunningham, G. B. (2015). Examining Social Capital Development among Volunteers of a Multinational Sport-for-Development Event. *Journal of Sport Management*, 29 (1), 27-41. doi: 10.1123/JSM.2013-0325.
- Welty Peachey, J. & Burton, L. (2017). Servant Leadership in Sport for Development and Peace: A Way Forward. *Quest*, 69 (1), 125-139. doi: 10.1080/00336297.2016.1165123.
- Welty Peachey, J. & Cohen, A. (2015). Reflections from scholars on barriers and strategies in sport-for-development research. *Journal of Sport for Development*, 3 (4), 16-27.
- Welty Peachey, J. & Cohen, A. (2016). Research Partnerships in Sport for Development and Peace: Challenges, Barriers, and Strategies. *Journal of Sport Management*, 30 (3), 282-297. doi: 10.1123/jsm.2014-0288.
- Welty Peachey, J., Lyras, A., Cohen, A., Bruening, J. E. & Cunningham, G. B. (2014). Exploring the Motives and Retention Factors of Sport-For-Development Volunteers. *Nonprofit and Voluntary Sector Quarterly*, 43 (6), 1052-1069. doi: 10.1177/0899764013501579.
- Welty Peachey, J., Schlenker, N. & Hill, P. (2019). Sport-for-development: A comprehensive analysis of theoretical and conceptual advancements. *Sport Management Review*, in press. doi: 10.1016/j.smr.2019.11.002.
- Wessels, J. (2020). Verankerung von Evaluation in der Politik. Ergebnisse einer Befragung der Arbeitskreise der DeGEval – Gesellschaft für Evaluation e.V. *Zeitschrift für Evaluation (ZfEv)*, 2020 (02), 349-363. doi: 10.31244/zfe.2020.02.10.
- White, H. & Bamberger, M. (2008). Introduction. Impact Evaluation in Official Development Agencies. *IDS Bulletin*, 39 (1), 1-11. doi: 10.1111/j.1759-5436.2008.tb00428.x.
- White, H. & Barbu, A. (2006). *Impact evaluation - the experience of the independent evaluation group of the World Bank* (Independent Evaluation Group, World Bank, Hrsg.). Zugriff am 04. Dezember 2019 unter <http://documents.worldbank.org/curated/en/475491468138595632/pdf/382680Impact1e10experience01PUBLIC1.pdf>
- Whitley, M. A., Farrell, K., Wolff, E. A. & Hillyer, S. J. (2019b). Sport for development and peace: Surveying actors in the field. *Journal of Sport for Development*, 7 (12), 1-15.
- Whitley, M. A., Fraser, A., Dudfield, O., Yarrow, P. & van der Merwe, N. (2020). Insights on the funding landscape for monitoring, evaluation and research in sport for development. *Journal of Sport for Development*, 8 (14), 21-35.
- Whitley, M. A., Massey, W. V., Camiré, M., Blom, L. C., Chawansky, M., Forde, S. et al. (2019a). A systematic review of sport for development interventions across six

- global cities. *Sport Management Review*, 22 (2), 181-193. doi: 10.1016/j.smr.2018.06.013.
- Widmer, T. (2000). Qualität der Evaluation - Wenn Wissenschaft zur praktischen Kunst wird. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 77-102). Opladen: Leske + Budrich.
- Widmer, T., Beywl, W. & Fabian, C. (Hrsg.). (2009). *Evaluation. Ein systematisches Handbuch* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Win, E. (2004). "If it doesn't fit on the blue square it's out!" An open letter to my donor friend. In L. C. Groves & R. B. Hinton (Eds.), *Inclusive aid. Changing power and relationships in international development* (pp. 123-127). London: Earthscan.
- Wirsich, N. (2018). *The GIZ Project "Yalla Riyada Bus". Exchange, Education and Conflict Transformation through Sport for Development*. Zugriff am 22. Oktober 2020 unter <https://www.sport-for-development.com/imglib/downloads/giz2018-en-Factsheet-Yalla%20Riyada%20Bus.pdf>
- Wnendt, W. (2011). Internationale Sportförderung aus Sicht des Auswärtigen Amtes. In K. Petry, M. Groll & W. Tokarski (Hrsg.), *Sport und internationale Entwicklungszusammenarbeit. Theorie- und Praxisfelder* (S. 107-118). Köln: Sportverlag Strauß.
- Wollmann, H. (2000). Evaluierung und Evaluationsforschung von Verwaltungspolitik und Modernisierung - zwischen Analysepotential und -defizit. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 195-232). Opladen: Leske + Budrich.
- Wollmann, H. (2017). Entwicklungslinien von Evaluation und Evaluationsforschung in Deutschland. Zwischen Vorgeschichte, Aufbruch und Konsolidierung. *Zeitschrift für Evaluation (ZfEv)*, 16 (2), 33-53.
- Wood, B., Betts, J., Etta, F., Gayfer, J., Kabell, D., Ngwira, N. et al. (2011). *The Evaluation of the Paris Declaration. Final Report*. Kopenhagen: Danish Institute for International Studies. Zugriff am 25. November 2019 unter <http://pd-website.in-force.dk/content/pdf/PD-EN-web.pdf>
- Wopp, C. (1995). *Entwicklungen und Perspektiven des Freizeitsports* (Edition Sport & Wissenschaft, 20). Aachen: Meyer & Meyer.
- Wopp, C. & Dieckert, J. (2002). Sport und Freizeit - Freizeitsport. In J. Dieckert & C. Wopp (Hrsg.), *Handbuch Freizeitsport* (Beiträge zur Lehre und Forschung im Sport, 134, S. 11-21). Schorndorf: Hofmann.
- Wright, P. M., Jacobs, J. M., Ressler, J. D. & Jung, J. (2016). Teaching for transformative educational experience in a sport for development program. *Sport, Education and Society*, 21 (4), 531-548. doi: 10.1080/13573322.2016.1142433.
- Wydra, G. (2007). *Sportpädagogik zwischen schulischer Pflicht, Gesundheitsorientierung und Erlebnishunger. Skript zur Vorlesung Sportpädagogik*. Saarbrücken: Sportwissenschaftliches Institut der Universität des Saarlandes.
- Young, K. & Okada, C. (Eds.). (2014). *Sport, Social Development and Peace* (Research in the Sociology of Sport, v.8). Bradford: Emerald Group Publishing Limited.

- Zanotti, L. & Stephenson, M. (2019). SDP and social theory. In H. Collison, S. C. Darnell, R. Giulianotti & P. D. Howe (Eds.), *Routledge Handbook of Sport for Development and Peace* (pp. 165-174). Milton: Routledge.
- Zeuner, A. (2015). Zum Auftrag des Schulsports, Erziehung zum und durch Sport. *Sportunterricht*, 64 (3), 86-89.
- Zintl, M. (2006). Evaluierung der Entwicklungszusammenarbeit: BMZ. *Zeitschrift der Arbeitsgemeinschaft Entwicklungsethnologie*, 15 (1+2), 41-66.
- Zintl, M. (2009). Evaluierung in der deutschen Entwicklungszusammenarbeit. In T. Widmer, W. Beywl & C. Fabian (Hrsg.), *Evaluation. Ein systematisches Handbuch* (1. Aufl., S. 245-254). Wiesbaden: VS Verlag für Sozialwissenschaften.