

Christina Holewik  
Unter Mitarbeit von Andreas Detzel, Frieder Rubik

---

# Evaluation in Reallaboren

## Hintergrundpapier



# Impressum

**Autor/innen:**

Christina Holewik (ifeu)

Unter Mitarbeit von Andreas Detzel (ifeu), Frieder Rubik (IÖW)

**Projektleitung:**

Institut für ökologische Wirtschaftsforschung (IÖW)

Potsdamer Str. 105, 10785 Berlin

[www.ioew.de](http://www.ioew.de)

**Kooperationspartner:**

ifeu – Institut für Energie- und Umweltforschung Heidelberg gGmbH

Wilckensstraße 3, 69120 Heidelberg

[www.ifeu.de](http://www.ifeu.de)

Der vorliegende Beitrag entstand im Forschungsprojekt „Innoredux – Geschäftsmodelle zur Reduktion von Plastikmüll entlang der Wertschöpfungskette: Wege zu innovativen Trends im Handel“. Das Projekt ist Teil des Forschungsschwerpunkts „Plastik in der Umwelt – Quellen, Senken, Lösungsansätze“ und wird gefördert vom Bundesministerium für Bildung und Forschung (BMBF).

Förderkennzeichen 01UP1804A

**Zitiervorschlag:**

Christina Holewik (2022): Evaluation in Reallaboren - Hintergrundpapier, Arbeitsbericht des Forschungsprojekts Innoredux.

Mehr Informationen zum Projekt: [www.plastik-reduzieren.de](http://www.plastik-reduzieren.de)

Heidelberg, Juli 2022

## Zusammenfassung

Für die Auseinandersetzung mit den vielfältigen aktuellen gesellschaftlichen Herausforderungen bieten Reallabore ein Forschungsformat, welches einen aktiven, transdisziplinären und transformativen Ansatz verfolgt. Die Durchführung von Evaluationen ist dabei relevant, da dadurch Wissen über die Wirkung oder Qualität von Maßnahmen generiert wird und somit Lern- und Transformationsprozesse weiterentwickelt werden können. In diesem auf einer Bachelor-Arbeit basierenden Hintergrundpapier werden die zum Zeitpunkt der Bachelorarbeit öffentlich zugänglichen Informationen über die Evaluation von Reallaboren systematisiert und damit ein Überblick zum aktuellen Stand gewonnen. Mittels einer tiefgehenden qualitativen Literaturrecherche wird zudem betrachtet, wie sich die Evaluationspraxis in den Baden-Württemberg (BaWü)-Labs konkret ausgestaltet. Dabei werden die Aspekte Evaluationsgegenstand, Zweck der Evaluation, Kriterienauswahl, Methodenauswahl, Evaluierende sowie Evaluationszeitpunkt betrachtet. Aus der Analyse ergibt sich, dass in den BaWü-Labs nicht nur extern durch das Evaluationsgremium evaluiert wurde, sondern ebenso intern in den Reallaboren selbst. In diesen internen Evaluationen der BaWü-Labs wurden vor allem umgesetzte Maßnahmen bzw. Realexperimente evaluiert, wozu teilweise Erfolgskriterien genutzt und Methoden unterschiedlicher Partizipationsgrade gewählt wurden. Des Weiteren erfolgte die Evaluation oftmals in mehrmaligen Evaluationsdurchführungen. Ausgehend von diesen strukturiert gesammelten und aufbereiteten Informationen werden praxisorientierte Empfehlungen für die Konzeption und Umsetzung der Evaluation des Reallabors im Projekt Innoredux abgeleitet und diskutiert. Damit bietet dieses Hintergrundpapier eine Zusammenführung des Wissens über mögliche Evaluationsansätze, einen Einblick in die Evaluationspraxis der BaWü-Labs sowie Empfehlungen für einen konkreten intraprojekt dynamischen Reallaboransatz. Hiermit soll sowohl einem weiteren Publikum an Evaluator\*innen und Forschenden ein Überblick über bisherige Ansätze und Praktiken als auch den Forschenden des Innoredux-Projekts eine erste Orientierung für die Evaluation ihres Reallabors gegeben werden.

## Abstract

Real-world-laboratories offer a research format for dealing with the diverse challenges society faces today, and follows an active, transdisciplinary, and transformative approach. The performance of evaluations is relevant in this context, as they generate knowledge about the impact or quality of measures and can thus further develop learning and transformation processes. In this background paper, which is based on a bachelor thesis, the information available at the time of the bachelor thesis with regard to the evaluation of real-world-laboratories was systematised and thus an overview over the current status was gained. By means of a more in-depth qualitative literature research, it was also examined how the evaluation practice in the Baden-Württemberg (BaWü)-Labs is designed in particular. The following aspects were considered: object of evaluation, purpose of evaluation, selection of criteria, selection of methods, evaluators, and time of evaluation. The analysis showed that evaluation in the BaWü-Labs was not only carried out externally by the evaluation committee, but also internally in the real-world-laboratories themselves. In these internal evaluations of the BaWü-Labs mainly implemented measures or real-world-experiments were evaluated, for which success criteria were used in some cases and methods with different degrees of participation were chosen. In addition, the evaluation was often carried out several times. Based on these structured and refined information, practice-oriented recommendations for the evaluation of the real-world-laboratory in the Innoredux project were derived and discussed. Thus, this background paper offers a consolidation of the knowledge about possible evaluation concepts, an insight into the evaluation practice of the BaWü-Labs as well as recommendations for a specific intra-project dynamic real-world-laboratory approach. The aim is to give an overview of previous approaches and practices to a wider audience of evaluators and to give researchers of the Innoredux project a first orientation for the evaluation of their real-world-laboratory.

# Inhaltsverzeichnis

<b>1</b>	<b>Vorwort .....</b>	<b>7</b>
<b>2</b>	<b>Einleitung.....</b>	<b>8</b>
<b>3</b>	<b>Konzeptionelle und begriffliche Aspekte von Reallaboren .....</b>	<b>9</b>
3.1	Beitrag zur Transformation.....	10
3.2	Experimente als Kernforschungsmethode .....	11
3.3	Transdisziplinarität als Forschungsmodus.....	12
3.4	Langzeitorientierung, Skalierbarkeit und Übertragbarkeit der Ergebnisse .....	12
3.5	Lernen und Reflexion .....	13
<b>4</b>	<b>Konzeptionelle und begriffliche Aspekte von Evaluation .....</b>	<b>14</b>
4.1	Gegenstand der Evaluation.....	15
4.2	Zweck der Evaluation .....	15
4.3	Evaluationskriterien.....	15
4.4	Evaluator*innen .....	16
4.5	Evaluationsmethoden.....	17
<b>5</b>	<b>Evaluation in Reallaboren .....</b>	<b>17</b>
5.1	Stand der Forschung.....	17
5.1.1	Co-Evaluation als Teil eines zirkulären Reallabormodells .....	18
5.1.1.1	Erfassung konkreter Ergebnisse .....	19
5.1.1.2	Co-Interpretation .....	22
5.1.1.3	Transfer in Wissenschaft und Praxis .....	23
5.1.2	Begleitende externe Evaluation durch ein Evaluationsgremium .....	24
5.1.3	Verschiedene Maßstäbe der Evaluation.....	26
5.2	Untersuchung der BaWü-Labs.....	28
5.2.1	Methodik .....	28
5.2.2	Ergebnisse .....	30
5.3	Was lernen wir daraus? .....	42
<b>6</b>	<b>Übertrag Innoredux.....</b>	<b>43</b>
6.1	Einbettung von Reallabor und Evaluation in das Gesamtprojekt.....	43
6.2	Akteurskonstellation .....	44
6.3	Konzeptionelle Anknüpfungspunkte.....	45
<b>7</b>	<b>Fazit.....</b>	<b>50</b>
<b>8</b>	<b>Literaturverzeichnis.....</b>	<b>51</b>
<b>9</b>	<b>Anhang .....</b>	<b>54</b>

## Abbildungsverzeichnis

Abb. 1: Grundlegende Aspekte der Evaluation.....	14
Abb. 2: Das zirkuläre Reallabormodell.....	19
Abb. 3: Erfassung konkreter Ergebnisse im Kontext der Co-Evaluation.....	20
Abb. 4: Modell der linearen Wirkungskette.....	21
Abb. 5: Formen produktiver Interaktionen. ....	22
Abb. 6: Transfer in Wissenschaft und Praxis. ....	24
Abb. 7: Einteilung der Evaluationsgegenstände nach Maßstab und Evaluationsart.....	26
Abb. 8: Projektablauf von Innoredux. ....	44
Abb. 9: Akteurskonstellation im Projekt Innoredux.....	45

## Tabellenverzeichnis

Tab. 1: Realexperiment und naturwissenschaftliches Experiment im Vergleich.....	11
Tab. 2: Einteilung der Evaluation nach Phasen und den dazugehörigen Evaluationsarten und - kriterien.....	16
Tab. 3: Übersicht über die gefundenen Informationsquellen zu Evaluation in den BaWü-Labs.....	29
Tab. 4: Kontext der Reallabore, zu welchen Informationen über die Evaluation vorliegen. ....	32
Tab. 5: Evaluationspraxis der betrachteten Reallabore. ....	34
Tab. 6: Sortierung der verwendeten Methoden in den betrachteten Reallaboren nach ihrem Partizipationsgrad.....	41
Tab. 7: Leitfragen und komprimierte Übersicht der Empfehlungen und Leitgedanken.....	46
Tab. 8: Abschließende Zusammenschau mit möglichen Auswahloptionen zur Entwicklung des Evaluationsdesigns.....	49
Tab. 9: Reflexionsschema. ....	54
Tab. 10: Schema für die Evaluation von Realexperimenten.....	54
Tab. 11: Gütekriterien für Methoden der partizipativen Wissenserzeugung in der transdisziplinären Forschung.....	60

# 1 Vorwort

Das Forschungsprojekt „Geschäftsmodelle zur Reduktion von Plastikmüll entlang der Wertschöpfungskette: Wege zu innovativen Trends im Handel“ (Innoredux) untersucht Geschäftsmodellinnovationen im Handel zur Reduktion des Plastikmüllaufkommens entlang der Wertschöpfungskette. Innoredux wird vom Bundesministerium für Bildung und Forschung (BMBF) im Forschungsschwerpunkt „Plastik in der Umwelt - Quellen, Senken, Lösungsansätze“ gefördert, Bearbeitungszeitraum ist von Februar 2019 bis Januar 2022. Ziel von Innoredux ist es, in einem Reallaborforschungsansatz gemeinsam mit Partnern aus der unternehmerischen und kommunalen Praxis sowie mit Verbänden eine praktische Umsetzung von Verpackungslösungen im Einzelhandel zu entwickeln. Betrachtet werden sowohl der stationäre Handel als auch der Online-Versandhandel, wobei der Fokus auf den Verpackungen von Produkten aus vier Warengruppen liegt: Lebensmittel, Textilien, Bürobedarf sowie Kosmetika, Hygiene-, Wasch-, Putz- und Reinigungsmittel (WPR).

Die Strategien zur Reduktion von (Kunststoff-)verpackungen werden methodisch aus einer Geschäftsmodellperspektive heraus konzipiert und im Zuge eines in der Stadt Heideberg angesetzten Reallabors erprobt. Innoredux gliedert sich in vier Arbeitspakete:

- Das erste Arbeitspaket typologisiert plastikmüllvermeidende und -reduzierende Geschäftsmodelle. Dabei werden sowohl innovative Verpackungslösungen als auch Geschäftsmodellinnovationen betrachtet.
- Im zweiten Arbeitspaket werden in Zusammenarbeit mit den Praxispartnern des Vorhabens instruktive Beispiele mit Blick auf ökologische, ökonomische und soziale Wirkungen untersucht, interne und externe Einflussfaktoren ermittelt sowie Ansatzpunkte für kommunales bzw. regionales Handeln analysiert.
- Das darauffolgende dritte Arbeitspaket schafft in Form eines Reallabors in einem geographisch und zeitlich abgegrenzten Raum einen realen Anwendungskontext, in dem Lösungen aus dem zweiten Arbeitspaket erprobt werden können.
- Schließlich werden im finalen vierten Arbeitspaket die gewonnen konzeptionellen und empirischen Ergebnisse ausgewertet und zu Strategien in Form von Handreichungen für Kommunen und Unternehmenschecklisten verdichtet.

Dieses Hintergrundpapier basiert auf einer Bachelorarbeit, die in Vorbereitung des Reallabors in Arbeitspaket 3 verfasst wurde. Konkret befasst sich das Papier damit, die vorliegenden Informationen hinsichtlich der Evaluation von Reallaboren zu systematisieren und daraus Empfehlungen für eine mögliche Gestaltung der Evaluation des Reallabors in Innoredux abzuleiten.

## 2 Einleitung

Wöchentliche Protestaktionen, Social-Media-Posts von Plastikbergen, Nachhaltigkeitsinitiativen im Handel – Wenn auch die Ausführungen mal mehr und mal weniger Aufmerksamkeit erregen, so wird in der öffentlichen Wahrnehmung doch eine Aussage immer lauter: Es soll sich etwas verändern.

Eine Antwort auf die Forderung nach gesellschaftlicher Transformation ist die Entwicklung des Forschungsformats „Reallabor“, da dieses auf realweltliche Probleme verschiedener Art eingeht (Wagner und Grunwald 2019, S. 261). Das heißt, es werden Herausforderungen betrachtet, welche nicht nur in akademischen Kreisen Aufmerksamkeit erlangen, sondern gerade von gesellschaftlicher Relevanz sind. Nahezu folgerichtig scheint es, dass das Format einen transdisziplinären Ansatz verfolgt, das heißt eine Zusammenarbeit zwischen wissenschaftlichen und außerwissenschaftlichen Akteuren (ebd., S. 261).

Ein Aspekt, der in diesem Forschungsfeld bisher jedoch weniger ausführlich behandelt wurde, ist die Evaluation der Reallabore. Zwar findet der Begriff immer wieder Erwähnung und es finden sich vereinzelt Publikationen und konzeptionelle Unternehmungen, eine Übersicht der vorhandenen Ansätze und Methoden ist bisher jedoch nach Kenntnis der Autor\*innen noch nicht vorhanden. Dabei kann Evaluation individuelles und organisationales Lernen unterstützen und so helfen soziale Wandlungsprozesse voranzutreiben (Luederitz et al. 2017, S. 62). Mit anderen Worten: Eine genauere Betrachtung der Evaluation in Reallaboren kann dem transformativen Anspruch des Forschungsformats durchaus dienlich sein. Damit eine Evaluation möglichst optimal an eine Untersuchungsfrage angepasst werden kann, ist es wichtig, dass Evaluator\*innen die Bandbreite an möglichen Untersuchungsdesigns und den dazugehörigen Methoden kennen (Stockmann und Meyer 2014, S. 18). Durch die derzeitige Verstreuung des Wissens, war es bisher jedoch nicht ohne weiteres möglich diese Bandbreite an Evaluationsansätzen in Reallaboren ausfindig zu machen.

Ein Ziel dieser Arbeit soll es deshalb sein eine erste Systematisierung der existierenden Ansätze zu unternehmen. Dies erfolgt durch die Sichtung der vorhandenen Literatur. Dabei wird zunächst auf den allgemeinen Stand der Forschung hinsichtlich der Evaluation von Reallaboren eingegangen. Anschließend wird mittels einer systematischen Literaturrecherche die Evaluationspraxis in den sog. Baden-Württemberg-Labs (BaWü-Labs) genauer untersucht.

In einem nächsten Schritt sollen die vorab gewonnen, theoretischen Erkenntnisse zur Entwicklung von Empfehlungen, hinsichtlich evaluationsbezogener Aspekte, für ein intraprojekt-dynamisch<sup>1</sup> entstehendes Reallabor im Projekt Innoredux genutzt werden. Hierbei soll das zuvor abstrahierte Wissen für eine Praxisnutzung vorbereitet, jedoch noch keine konzeptionelle Arbeit geleistet werden. Vielmehr soll mittels der hier dargelegten Empfehlungen eine erste Praxisorientierung gegeben werden.

Zusammengefasst prägen die folgenden Forschungsfragen die vorliegende Arbeit:

- Wie wurden Reallabore bisher evaluiert?

---

<sup>1</sup> Innerhalb des Projekts entstehendes Reallabor, detaillierte Erklärung zur Begriffswahl siehe Kapitel 6.1

- Welche Methodik könnte für die prozessbegleitende Evaluation eines intraprojekt-dynamisch entstehenden Reallabors Anwendung finden?

Im folgenden Kapitel 3 werden dazu zunächst die Kerncharakteristika und die besondere Akteurskonstellation in Reallaboren erläutert. Im Anschluss werden konzeptionelle und begriffliche Aspekte der Evaluation thematisiert (Kapitel 4). Auf diesen beiden Kapiteln aufbauend folgt eine detailliertere Ausführung zu den bisherigen Evaluationsansätzen in Reallaboren (Kapitel 5). Konkret sollen dabei, nach einer allgemeinen Begriffsklärung, die zwei Ansätze der Co-Evaluation sowie des externen Evaluationsgremiums erklärt werden. Weitere Ansätze werden nach der Maßstabsebene des Evaluationsgegenstands systematisiert (Kapitel 5.1). Es folgt eine Erklärung der methodischen Herangehensweise der durchgeführten Literaturrecherche. Mit der Aufschlüsselung der Evaluationspraxis der einzelnen BaWü-Labs wird der Ergebnis-Teil der Arbeit gestaltet (Kapitel 5.2). Daran anknüpfend werden in Kapitel 6 Empfehlungen für die Evaluation im Reallabor im Projekt Innoredux diskutiert. Abschließend wird ein Fazit gezogen (Kapitel 7).

### 3 Konzeptionelle und begriffliche Aspekte von Reallaboren

In diesem Kapitel soll genauer auf das Forschungsformat Reallabor eingegangen werden. Denn um sich der Frage anzunähern, wie Reallabore bisher evaluiert wurden, ist es wichtig zu verstehen, was ein Reallabor eigentlich charakterisiert. So lassen sich die besonderen Anforderungen bzw. die spezifischen Rahmenbedingungen für die Evaluation später besser einordnen.

Reallabore sind noch ein relativ junger Forschungsansatz (Rose et al. 2019, S. 1). Nach einer Ausschreibung des Ministeriums für Wissenschaft, Forschung und Kunst (MWK) im Jahr 2013 (Wagner 2018) hat Baden-Württemberg als erstes Bundesland in Deutschland ein Förderprogramm für Reallabore ins Leben gerufen (Parodi et al. 2018, S. 1). So konnten von 2015 bis 2019 14 Reallabore, die so genannten BaWü-Labs, in zwei zeitversetzten Förderlinien aufgebaut werden (Defila und Di Giulio 2019b, S. 4). Die Förderung umfasste auch zwei Begleitforschungsprojekte (ebd., S. 4) und ein vom MWK beauftragtes Evaluationsgremium (Zimpelmann 2018, S. 321). Seitdem hat das Forschungsformat auch bei anderen Fördergebern zunehmend Anklang gefunden, wie sich an weiteren Ausschreibungen zeigt (Defila und Di Giulio 2019b, S. 1). Seit April 2019 gibt es zudem das „Netzwerk Reallabor der Nachhaltigkeit“ (Hartlieb 2019). Gegründet durch das Karlsruher Institut für Technologie (KIT), das Wuppertaler Institut für Klima, Umwelt und Energie (WI), die Leuphana Universität Lüneburg und das Ecological Research Network (Ecor-net) bietet dieses junge Netzwerk eine Kommunikationsplattform im Kontext von Reallaboren (ebd.). An diesen Entwicklungen zeichnet sich ab, dass Reallabore zunehmend an Bedeutung und Popularität gewinnen.

Wichtig zu erwähnen ist, dass Reallabore nur eine mögliche Variante unter anderen realweltlich verankerten Forschungskonzepten sind (Rose et al. 2019, S. 3). So nennen Rose et al. (2019, S. 3) als weitere nationale und internationale Konzepte beispielsweise noch „(sustainable) living labs“, „urban transition labs“, „T-labs“ oder „transition management“. Der Diskurs um Reallabore spielt sich vor allem in Deutschland ab, wobei Ähnlichkeiten zu den

Entwicklungen des „transition managements“ und den „urban transition labs“ in den Niederlanden festzustellen sind (Rose et al. 2019, S. 3). Integriert werden aber auch Ansätze der transdisziplinären Nachhaltigkeitsforschung sowie der Aktionsforschung (ebd., S. 3).

Nachdem aufgrund des jungen Alters des Forschungsformats noch unterschiedliche Konzeptionsvorschläge von Reallaboren vorliegen, haben Schöpke et al. (2018, S. 86) fünf Kerncharakteristiken herausgearbeitet:

- Beitrag zur Transformation
- Experimente als Forschungsmethode
- Transdisziplinarität als Forschungsmodus
- Langzeitorientierung, Skalierbarkeit und Übertragbarkeit der Ergebnisse
- Lernen und Reflexion

### 3.1 Beitrag zur Transformation

Das erste Kerncharakteristikum von Reallaboren nach Schöpke et al. (2018, S. 86) ist der Beitrag zur Transformation. Diese wird implizit definiert als Beitrag zur Nachhaltigkeitstransformation: „Research aiming to contribute to sustainability transformation“ (ebd., S. 86). Reallabore kontribuieren hierbei sowohl zur Transformation, als auch zur Transformationsforschung (ebd., S. 86). Zu Ersterem tragen sie durch die Durchführung von Realexperimenten bei, welche potentielle Lösungen für gesellschaftliche Herausforderungen testen (ebd., S. 86). Zu Zweiterem dadurch, dass sie in diesem Prozess Wissen generieren und Hinweise über die Robustheit, Skalierbarkeit und Übertragbarkeit der Lösungen geben (ebd., S. 86f.). Durch das aktive Eingreifen in soziale Wandlungsprozesse, ist das Reallabor in einen politischen und normativen Rahmen eingebettet (ebd., S. 87). Wagner und Grunwald (2019, S. 261) schreiben weiter, dass Reallabore wissenschaftliche Erkenntnisse über gesellschaftliche Wandelprozesse zu einer Kultur der Nachhaltigkeit erbringen und diese Prozesse ebenso aktiv anregen sollen. Das heißt das Forschungsformat ist a priori nicht neutral oder deskriptiv, sondern hat nach diesem Verständnis einen wertenden Bezug zu Nachhaltigkeit bzw. diese als Norm.

Defila und Di Giulio (2018b, S. 11) sowie Borner und Kraft (2018, S. 3) weisen jedoch darauf hin, dass Reallabore nicht zwingend dem Ziel einer nachhaltigen Entwicklung verpflichtet sein müssen. Dennoch unterliegen die Ziele eines Reallabors aufgrund der aktiv anvisierten gesellschaftlichen Veränderung einer besonderen Begründungs- und Rechenschaftspflicht (Defila und Di Giulio 2018b, S. 11f.). Hinzu kommt die oftmalige Verwendung öffentlicher Mittel, welche ebenfalls mit einer Verpflichtung gegenüber ebenjener Öffentlichkeit einhergeht (ebd., S. 12). Die Autor\*innen schlagen deshalb vor, dass sich die Ziele eines Reallabors respektive der transformativen Forschung im Allgemeinen daran orientieren sollen, dass sie „gesellschaftlich legitimiert, ethisch gut begründet und gemeinwohlorientiert sind“ (ebd., S. 12). Damit soll eine Willkür der Zielausrichtung ausgeschlossen werden, ohne aber eine Orientierung an Nachhaltigkeit vorauszusetzen (ebd., S. 12). Die Gutachtergruppe der BaWü-Labs distanzieren sich jedoch von dem von Defila und Di Giulio (2018b) entwickelten Konzept und empfehlen „bei der Weiterentwicklung des Förderformats Reallabore den Fokus auf die Nachhaltigkeit zu reaktivieren und zu präzisieren“ (Zimpelmann 2018, S. 322).

Inwiefern Nachhaltigkeit also ein zwingendes Kriterium sein sollte, ist demnach noch umstritten, nichtsdestotrotz ist der Transformationsbegriff ein Kerncharakteristikum von Reallaboren.

## 3.2 Experimente als Kernforschungsmethode

So wie in einem naturwissenschaftlich definierten Labor Experimente zu erwarten sind, werden in Reallaboren Realexperimente durchgeführt. Dabei bietet sich mittels der transformativen Rahmenbedingungen von Reallaboren die Möglichkeit, zuvor erarbeitete Lösungsoptionen zu testen und ihren Nutzen zu erproben (Borner und Kraft 2018, S. 3f.). Die Grenzen der Realexperimente werden durch den zeitlichen und räumlichen Rahmen des Reallabors bestimmt (ebd., S. 4). Realexperimente stellen „realweltliche Interventionen“ dar (Rose et al. 2019, S. 5). Sie sind also Eingriffe in ein bestehendes System (Arnold und Piontek 2018, S. 147). Anders als in einem naturwissenschaftlichen Experiment, mit isoliertem Versuchsaufbau, finden diese Interventionen allerdings im öffentlichen Raum statt (ebd., S. 147). Durch die daraus entstehende Abhängigkeit vom jeweiligen sozialen Kontext, mit seinen stetigen gesellschaftlichen Wandlungsprozessen, fällt die Möglichkeit kontrollierbarer Laborbedingungen jedoch weg (ebd., S. 147). Dennoch können hierbei mögliche (ggf. innovative) Lösungsstrategien erprobt und auf ihre Anwendbarkeit getestet werden. In Tab. 1 sind die Charakteristika bzw. Unterschiede von Realexperimenten im Vergleich zum naturwissenschaftlichen Experimentalansatz, wie sie Arnold und Piontek (2018, S. 147) aufgearbeitet haben, tabellarisch dargestellt.

**Tab. 1: Realexperiment und naturwissenschaftliches Experiment im Vergleich.**

Dargestellt ist eine Gegenüberstellung der einzelnen Merkmale der beiden Experimentbegriffe, um sie einerseits zu definieren und andererseits voneinander abzugrenzen (Arnold und Piontek 2018, S. 147).

	Realexperiment	Experiment
Primäres Ziel	Schaffung von Wissen zur Lösung gesellschaftlicher Problemlagen	Schaffung von Wissen zur Lösung wissenschaftlicher Probleme
Dokumentation	Dokumentation dem Thema, den Akteuren und der Intervention entsprechend und wie unter den jeweiligen Bedingungen umsetzbar	Umfassende Dokumentation nach wissenschaftlichen Standards
Experimentelles Umfeld	Durchführung in einem lebensweltlichen ‚Raum‘	Durchführung im isolierten Labor
Durchführende	Alle Akteure eines Reallabors und Angehörige des jeweiligen lebensweltlichen ‚Raums‘	Wissenschaftliches Personal
Wiederholbarkeit	Kann aufgrund seines Kontextes (beständiger gesellschaftlicher Wandel,	Muss zur wissenschaftlichen Validität wiederholbar sein

	keine kontrolliert herstellbaren Laborbedingungen) nicht exakt wiederholt werden	
Reliabilität	Unkontrollierbare Bedingungen des Experimentes lassen kaum Aussagen zur Verlässlichkeit der Messungen zu	Hohes Maß der Verlässlichkeit der Messungen notwendig (Stabilität, Konsistenz, Äquivalenz)

### 3.3 Transdisziplinarität als Forschungsmodus

Der Forschungsmodus von Reallaboren ist grundsätzlich transdisziplinär (Wagner und Grunwald 2019, S. 261). Das heißt, das jeweilige Problemfeld wird kollaborativ von wissenschaftlichen und außerwissenschaftlichen Akteuren bearbeitet (ebd., S. 261). Dadurch sollen „Wissensaustausch und Wissensintegration auf Augenhöhe“ stattfinden (Borner und Kraft 2018, S. 4). Dabei ist entscheidend, dass es sich nicht nur um eine Konsolidierung der außerwissenschaftlichen Akteure, sondern um eine tatsächliche Zusammenarbeit handelt (Wanner et al. 2018, S. 101). Wanner et al. (2018, S. 101) sprechen in diesem Kontext von Co-Leadership der wissenschaftlichen Akteure mit den wichtigsten Praxisakteuren. Es geht also um eine gemeinsame Projekt(durch-)führung, welche Wissenschafts- und Praxisgrenzen überschreitet.

### 3.4 Langzeitorientierung, Skalierbarkeit und Übertragbarkeit der Ergebnisse

Wie oben bereits erwähnt, sollen Reallabore gesellschaftliche Transformationsprozesse anregen (Schäpke et al. 2018, S. 87). Für solche langfristigen und großflächigen Veränderungen braucht es eine langzeitorientierte Forschungsperspektive, die über die Laufzeit des Reallabors hinaus sieht (ebd., S. 87). Hierfür ist es wichtig, Lösungen übertrag- und skalierbar zu machen (ebd., S. 87) oder in den Worten von Schäpke et al.: „Large-scale impacts do not necessarily depend on the long-term existence of RwLs<sup>2</sup>, but rather on the uptake of solutions“. Nachdem einfache Nachbildungen oder Wiederholungen durch die starke Kontextualität und Komplexität nicht ohne weiteres möglich sind, ist es wichtig, Wissen darüber zu generieren, wie Skalier- und Übertragbarkeit möglich sind. Dabei meint Übertragbarkeit, dass Wissen von einem Kontext in andere Kontexte eingebettet werden kann (ebd., S. 87). Dies ist beispielsweise durch Generalisierung von Erkenntnissen möglich oder durch Wissen über kontextuelle Faktoren (ebd., S. 87). Unter Skalierbarkeit wird hier verstanden, dass die Reichweite von entwickelten Lösungen vergrößert werden kann (ebd., S. 87). Dafür ist es notwendig, Wissen über skalierbare Faktoren zu erlangen (ebd., S. 87). Doch insbesondere die Transferier- und Skalierbarkeit stellen durch die starke Kontextgebundenheit der Reallabore eine große Herausforderung in der Reallaborforschung dar (ebd., S. 87).

<sup>2</sup> Realworld Labs

## 3.5 Lernen und Reflexion

Um Transformationsprozesse zu erreichen, müssen soziale Akteure neue Perspektiven, Fähigkeiten, Kompetenzen, Praktiken und Rollenverständnisse erlernen (Singer-Brodowski et al. 2018, S. 23). Deshalb ist es wichtig, dass transformative Ansätze, wie die Reallaborforschung, Lern- und Reflexionsprozesse etablieren, um ebendiese Entwicklungen möglich zu machen (ebd., S. 23). Die Reallabore selbst können hierbei nach Singer-Brodowski et al. (2018, S. 24) aus bildungswissenschaftlicher Perspektive als Lernumgebung dienen. Dabei ermöglichen sie Lernen auf den drei Ebenen der persönlichen Kompetenzentwicklung, sozialem Lernen sowie der inter- und transdisziplinärer Zusammenarbeit (ebd., S. 26).

Das individuelle Lernen, das heißt die Gewinnung von Kompetenzen zur Teilhabe an Transformationsprozessen, ergibt sich mittels erfahrungsbezogener Lernzyklen, basierend auf Wissensaustausch, Handlung und Reflektion (ebd., S. 24). Diese entstehen indem die Lernenden im Reallabor Teil des Lösungsprozesses werden, sowohl in der Wissenschaft als auch in der Praxis (ebd., S. 24). In diesen Lösungsprozessen wenden die Beteiligten zunächst ihr theoretisches Wissen auf ein konkretes Problem an (ebd., S. 24). Aus der Anwendung ergeben sich neue Erfahrungen, welche wiederum abstrahiert und damit in neues Wissen umgewandelt werden (ebd., S. 24). Aus diesem zweiseitigen Prozess der Wissensanwendung und Wissensgenerierung resultieren dann individuelle Kompetenzen dahingehend, wie die Lernenden selbst Teil von (Nachhaltigkeits-)Transformationen sein können (ebd., S. 24). Soziales Lernen bzw. informelle Lernprozesse aus gegenseitigem und gemeinsamen Lernen entstehen durch das Analysieren und Auseinandersetzen der diversen Stakeholder mit verschiedenen Perspektiven im Reallabor, in Folge der gemeinsamen Arbeit an einer konkreten Problemstellung (ebd., S. 25). Das Reallabor ist hierbei eine gemeinsame Lernumgebung, in welcher divergierende Perspektiven bzw. Vorannahmen, hinsichtlich der eigenen bedeutungsgebenden Denkweisen und Wahrnehmungen, analysiert und ausgehandelt werden können (ebd., S. 25). Dadurch ist es möglich den Blick über die bisherige Perspektiven zu weiten und einen höheren Grad der Reflexivität zu erreichen (ebd., S. 25). Dabei sind die Beteiligten Teil eines kollektiven Prozesses der Bedeutungszuweisung (Schäpke et al. 2018, S. 88). Durch die inter- und transdisziplinäre Zusammenarbeit werden des Weiteren ebenfalls Lernprozesse dahingehend angeregt, als dass ein Verständnis für verschiedene Communitys of Practice generiert werden muss, um miteinander kollaborieren zu können (Singer-Brodowski et al. 2018, S. 25). So müssen die diversen Akteursgruppen sich über die eigenen disziplinären Grenzen hinweg vertraut machen mit verschiedenen Herangehensweisen in ungewohnten Handlungsfeldern, um gegenseitiges Lernen voneinander zu ermöglichen (ebd., S. 26). Dabei werden verschiedene Formen von Wissen integriert sowie neue Rollen und Sprachen entwickelt werden, um miteinander kooperieren zu können (ebd., S. 25f.). Reflexivität ist in Reallaboren hierbei entscheidend, um verschiedene epistemische Kulturen, Werte und Ziele zu konfrontieren, in Beziehung zu setzen und zu integrieren (Schäpke et al. 2018, S. 88).

## 4 Konzeptionelle und begriffliche Aspekte von Evaluation

Zunächst soll der Begriff „Evaluation“ genauer eingegrenzt werden, da die alltagssprachliche Verwendung von der wissenschaftlichen Perspektive abweichen kann (Stockmann und Meyer 2014, S. 73). Dazu werden Kennzeichen und Begriffe erklärt, die für ein grundlegendes Verständnis relevant sind. Nachdem es sich bei Evaluation respektive der Evaluationsforschung um ein eigenes wissenschaftliches Feld handelt, kann hier allerdings nur ein Abriss der wichtigsten Termini gegeben werden. Mit Hilfe dieses Kapitels soll eine evaluations-theoretische Basis gelegt werden, um das Verständnis dieser Arbeit zu erleichtern. Für ein tieferes Verständnis sollte evaluationspezifische, weiterführende Literatur zu Rate gezogen werden.

Grundlegend kann Evaluation verstanden werden als ein „Instrument zur empirischen Generierung von Wissen [...], das mit einer Bewertung verknüpft wird, um zielgerichtete Entscheidungen zu treffen“ (Stockmann und Meyer 2014, S. 72). Es geht also darum, Wissen mittels empirischer Methodik zu sammeln, dieses anschließend zu bewerten, um auf Grundlage dessen eine fundierte Entscheidung treffen zu können.



**Abb. 1: Grundlegende Aspekte der Evaluation.** Bei der Evaluation wird empirisches Wissen gesammelt, anschließend bewertet, um darauf aufbauend zielgerichtete Entscheidungen zu treffen (eigene Darstellung nach Stockmann und Meyer 2014, S. 72).

Nach Stockmann und Meyer (2014, S. 74) weisen sich wissenschaftlich durchgeführte Evaluationen durch folgende sechs Kennzeichen aus:

1. Die Evaluation bezieht sich auf einen eindeutig definierten Gegenstand.
2. Es werden objektivierende, empirische Datenerhebungsmethoden für die Informationsgenerierung eingesetzt.
3. Die Bewertung geschieht auf Grundlage konkret festgelegter, transparenter Kriterien, die auf den zu evaluierenden Sachverhalt zugeschnitten sind.
4. Zur Bewertung werden systematisch vergleichende Verfahren eingesetzt.
5. Die Durchführung erfolgt durch befähigtes Personal, den sogenannten Evaluator\*innen oder Evaluierenden.
6. Das dahinterliegende Ziel der Evaluation ist das Treffen von Entscheidungen, welche sich auf den Evaluationsgegenstand beziehen.

Von diesen Kennzeichen ausgehend lassen sich Leitfragen ableiten, mit welchen sich eine professionell durchgeführte Evaluation auseinandersetzen sollte (ebd., S. 75). Konkret lauten diese „Was (welcher Gegenstand) wird wozu (zu welchem (Zweck) anhand welcher Kriterien von wem wie (mit welchen Methoden) evaluiert?“ (ebd., S. 75).

## 4.1 Gegenstand der Evaluation

Der Evaluationsgegenstand, auch Evaluationsobjekt oder Evaluandum genannt (Döring und Bortz 2016, S. 979), ist nach der DeGEval (2016, S. 67) „das, was im Rahmen der Evaluation untersucht und bewertet wird und worauf sich mögliche Konsequenzen in Folge der Evaluation beziehen“. Die Variationsbreite ist dabei sehr groß und kaum eingeschränkt (Stockmann und Meyer 2014, S. 75). Mit anderen Worten ist der Evaluationsgegenstand das Referenzobjekt, worauf sich die Evaluation bezieht und der Auswahl dessen sind kaum Grenzen gesetzt.

## 4.2 Zweck der Evaluation

Eine Evaluation kann verschiedenen Zwecken bzw. Funktionen dienen. Döring und Bortz (2016, S. 987) unterscheiden zwischen fünf verschiedenen Funktionen, nämlich der Erkenntnisfunktion, der Lern- und Dialogfunktion, der Optimierungsfunktion, der Entscheidungsfunktion und der Legitimationsfunktion. Erkenntnisfunktion meint, dass wissenschaftliche Erkenntnisse über das Evaluationsobjekt gesammelt werden (ebd., S. 987). In Abhängigkeit davon, in welcher Form die gewonnenen Erkenntnisse verbreitet und vermittelt werden, kann dadurch das Wissen der Beteiligten, Betroffenen, der Fachöffentlichkeit und/oder der breiten Öffentlichkeit vergrößert werden (ebd., S. 987). Die Lern- und Dialogfunktion beschreibt, dass eine Evaluation dazu genutzt werden kann, einen konstruktiven Austausch über den Evaluationsgegenstand zu unterstützen (ebd., S. 987). Dabei ist es von großer Bedeutung, die relevanten Stakeholder in den Prozess zu involvieren und Moderationsstrukturen zu etablieren (ebd., S. 987). Die Evaluation kann jedoch auch dazu dienen, den Evaluationsgegenstand zielgerichtet zu verbessern. So sollen im Fall der Optimierungsfunktion Erkenntnisse dazu generiert werden, wie das Evaluationsobjekt verbessert oder weiterentwickelt werden kann (ebd., S. 987). Bei der Entscheidungsfunktion geht es darum durch die Evaluation eine Grundlage für Entscheidungen zu schaffen (ebd., S. 987). Dabei geht es jedoch nicht darum, dass die Evaluator\*innen die Entscheidung treffen, es sollen lediglich Bewertungsgrundlagen oder Praxisempfehlungen für die Verantwortlichen geschaffen werden (ebd., S. 987). Unter Legitimationsfunktion wird der Zweck einer Evaluation verstanden, der die Entwicklung und Durchführung einer Intervention kontrolliert und dadurch nach Außen legitimiert (ebd., S. 987).

Der Zweck ist vom Ziel des Evaluationsgegenstands zu unterscheiden. Denn das Ziel stellt die „intendierte Wirkung des Gegenstands“ dar, also das, was letztendlich konkret erreicht werden soll, z.B. mit einer konkreten Maßnahme oder einem ganzen Projekt (DeGEval 2016, S. 71). Der Zweck stellt dahingehend also eher eine übergeordnete Ebene dar, in die die Ziele eingebettet sind.

## 4.3 Evaluationskriterien

Ein Evaluationskriterium ist ein Merkmal des betrachteten Evaluationsgegenstandes (ebd., S. 68). Diese Kriterien sind notwendig für eine wissenschaftlich fundierte Bewertung des Evaluationsobjekts (Döring und Bortz 2016, S. 983). Entscheidend ist, dass sich die Evaluationskriterien empirisch erfassen lassen und inhaltlich zur Lösung des Evaluationsproblems

beitragen (Döring und Bortz 2016, S. 983). Wie in Tab. 2 dargestellt, hängt die Kriterienauswahl auch davon ab, in welcher Phase evaluiert werden soll bzw. um welche Art der Evaluation es sich handelt.

**Tab. 2: Einteilung der Evaluation nach Phasen und den dazugehörigen Evaluationsarten und -kriterien.** Je nach Phase der Maßnahmenumsetzung hat die Evaluation eine andere Aufgabe, woraus sich unterschiedliche Arten der Evaluation und der benötigten Kriterien ergeben. Hier erfolgt eine dreifache Gliederung von der Konzeption einer Maßnahme bis zu ihrem Abschluss oder darüber hinaus (eigene Darstellung nach Döring und Bortz 2016, S.984).

Phase der Maßnahmenumsetzung / Aufgabe der Evaluation	Art der Evaluation	Art der Kriterien
Für die Maßnahme liegt ein Konzept vor. Die Evaluationsstudie soll im Vorfeld der potenziellen Realisierung die Umsetzbarkeit in der Praxis evaluieren.	Konzeptevaluation	Kriterien zur Bewertung des Konzepts einer Maßnahme
Die Maßnahme wird realisiert. Die Merkmale der Umsetzung sollen evaluiert werden, ggf. in Hinblick auf Verbesserungsbedarf.	Prozess- oder Implementierungsevaluation	Kriterien zur Bewertung der Durchführung einer Maßnahme
Die Maßnahme wird oder wurde durchgeführt. Die Outputs (unmittelbare Leistungen/Ergebnisse), Outcomes (kurzfristige Wirkungen) und/oder der Impact (langfristige, überindividuelle Folgen) soll evaluiert werden.	Ergebnis- bzw. Produktevaluation	Kriterien zur Bewertung der Ergebnisse einer Maßnahme

Die Auswahl der Kriterien kann auf verschiedene Weisen erfolgen. So gibt es teilweise Vorgaben bzw. verbindliche Evaluationskriterien, die vom Auftraggebenden festgelegt werden (Stockmann und Meyer 2014, S. 87). Die Festlegung der Evaluationskriterien kann jedoch auch wissens- oder erfahrungsbasiert durch die Evaluator\*innen selbst erfolgen (ebd., S. 87f.). Zudem gibt es die partizipative Kriterienauswahl, wobei z.B. Auftraggebende, Evaluator\*innen, Vertreter\*innen der Zielgruppe sowie weitere Stakeholder an der Kriterienfestlegung beteiligt sein können (ebd., S. 88).

## 4.4 Evaluator\*innen

Bei der Frage, wer die Evaluation durchführt, unterscheidet man zwischen externer und interner Evaluation. Bei der externen Evaluation sind die Evaluator\*innen nicht selbst Teil der Organisation, die für den Evaluationsgegenstand verantwortlich ist (DeGEval 2016, S. 66). Bei der internen Evaluation hingegen gehören Evaluierende und Evaluierete derselben Organisation an (ebd., S. 67). Bei der internen Evaluation lässt sich jedoch weiter nach Fremd- und Selbstevaluation unterscheiden. Eine Fremdevaluation, auch In-house-Evaluation genannt, besteht dann, wenn die Evaluator\*innen nicht in die Entwicklung, Umsetzung oder das Management des Evaluationsgegenstands involviert sind (ebd., S. 67). Das heißt die Rollen der Evaluierenden oder der Programmverantwortlichen sind klar getrennt (DeGEval

2016, S. 67). Die externe Evaluation ist dementsprechend auch automatisch eine Fremdevaluation. Bei der Selbstevaluation fallen die Rollen Evaluierende, Programmverantwortliche und teilweise auch Auftraggebende zusammen (ebd., S. 70), sprich die Evaluierenden sind auch in der operativen Durchführung beteiligt (Stockmann und Meyer 2014, S. 88).

## 4.5 Evaluationsmethoden

Für die Wahl der Evaluationsmethode(n) ist das zugrundeliegende Forschungsparadigma entscheidend (ebd., S. 92). Diese unterteilen Stockmann und Meyer (2014, S. 92) grob in zwei Hauptrichtungen: diejenigen, die einer kritisch-rationalen Forschungslogik folgen und diejenigen, die Realität als sozial konstruiert betrachten. Bei Ersterem wird Evaluation als angewandte Sozialforschung mit besonderen Forschungsbedingungen und spezifischem Erkenntnis- und Verwertungsinteresse verstanden, welche sich prinzipiell aller bekannten empirischen Forschungsmethoden bedienen kann (ebd., S. 92). Bei Zweiterem wird hingegen abgestritten, dass es eine objektiv zu beschreibende, reale Welt existiert, die sich mit Hilfe empirisch-wissenschaftlicher Verfahren beschreiben lässt (ebd., S. 92). Dadurch wird ein eher „[qualitatives] Denken“ gefordert, das verschiedene Perspektiven und Interpretationen der Realität erfasst (ebd., S. 92). Intersubjektive Nachprüfbarkeit soll dadurch erreicht werden, dass durch die Anwendung strenger methodologischer Regeln wahrnehmungsbeeinflussende Bedingungen weitestgehend unterdrückt werden (ebd., S. 92). Ähnlich zu dieser Unterscheidung schreiben Döring und Bortz (2016, S. 991), dass Evaluationen analog zu den wissenschaftstheoretischen Paradigmen in quantitativ-statistische, qualitativ-interpretative und Mixed-Methods-Evaluationen eingeteilt werden können. Die Unterscheidung nach quantitativen und qualitativen Studien spielt in der Praxis jedoch eine eher untergeordnete Rolle, da häufig sowohl quantitative als auch qualitative Daten als Methodenmix erhoben und ausgewertet werden können (ebd., S. 991). Wenn die qualitativen und quantitativen Daten nicht nur separat ausgewertet und interpretiert werden, sondern systematisch und wissenschaftstheoretisch reflektiert integriert werden, handelt es sich um einen Mixed-Methods-Ansatz (ebd., S. 991).

Mit anderen Worten ist es bei der Wahl der Evaluationsmethoden wichtig, sich grundlegend darüber Gedanken zu machen, welcher wissenschaftlichen Logik gefolgt werden soll. Demnach kann entschieden werden, welche Erhebungsmethoden für die Erhebung, Auswertung, Bewertung und Interpretation des Evaluationsgegenstandes adäquat sind.

# 5 Evaluation in Reallaboren

## 5.1 Stand der Forschung

Die Evaluation von Reallaboren ist bisher noch verhältnismäßig wenig erforscht. Dies liegt wohl vorrangig daran, dass Reallabore selbst noch relativ neu sind, im Vergleich zu anderen Forschungsformaten: „Bis dato liegen für das junge Format „Reallabor“ noch keine diskursiv und erfahrungsbasiert gefestigten Kriterien und Maßstäbe zur Evaluation vor“ (Parodi et al.

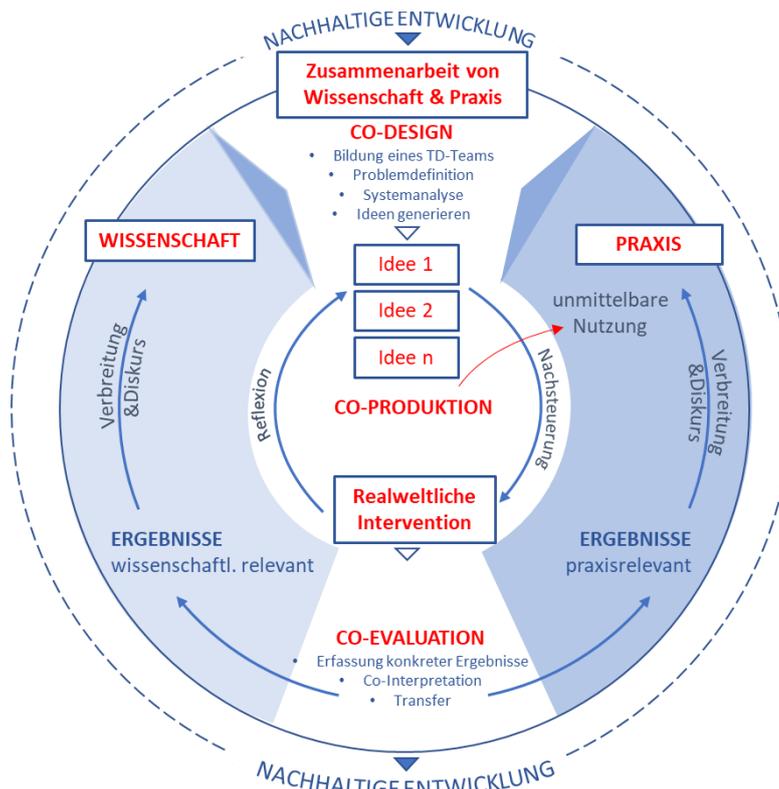
2018, S. 8). Nichtsdestoweniger findet Evaluation immer wieder Erwähnung in der reallaborbezogenen Literatur, wenn auch rein evaluationsspezifische Literatur eher selten ist (vgl. z.B. Luederitz et al. 2017, Geiger et al. 2017).

Insgesamt lassen sich aus der Literatur vor allem zwei Ansätze herauskristallisieren: Einerseits das Prinzip der Co-Evaluation, wie es im Kontext der Reallabore des Wuppertal Instituts konzeptualisiert wird (Rose et al. 2019; Wanner et al. 2018) und andererseits die begleitende Evaluation durch ein externes Evaluationsgremium, wie es von den BaWü-Labs empfohlen wird (Parodi et al. 2018). Auf beide Ansätze soll in den folgenden zwei Unterkapiteln nochmals näher eingegangen werden. Daran anschließend wird versucht, weitere Ansätze zu systematisieren und in ein verständliches, nach dem Maßstab der Evaluationsgegenstände geordnetes Schema zu sortieren.

### 5.1.1 Co-Evaluation als Teil eines zirkulären Reallabormodells

Im vom Wuppertal-Institut entwickelten Ablaufschema wird Evaluation als Teilprozess in der Gesamtkonzeption expliziert (siehe Abb. 4). Konkret wirkt sie gemeinsam mit Co-Design und Co-Produktion als Co-Evaluation im Zusammenspiel von Wissenschaft und Praxis mit (Wanner und Stelzer 2019, S. 3). Dabei ist das gesamte Reallabor in einen normativen Rahmen der Nachhaltigkeit eingebettet (Wanner et al. 2018, S. 102). Von der Evaluation gehen, wie in Abb. 2 veranschaulicht Erkenntnisse bzw. Impulse in Richtung der nachhaltigen Entwicklung heraus. Evaluation wird hier zudem als einer von zwei Lernzyklen verstanden (ebd., S. 102). So findet der erste Lernzyklus in Reflexions-Nachsteuerungs-Zyklen im Rahmen der Co-Produktion statt. Der zweite Lernzyklus findet durch den Transfer bzw. die Reintegration der (Evaluations-)Ergebnisse in Wissenschaft und Praxis statt (ebd., S. 102). Dadurch wird gegebenenfalls die Problemdefinition und das strukturelle Verständnis angepasst (ebd., S. 102).

Die Konzeption der Evaluation liegt in der Verantwortung der Wissenschaftler\*innen, ebenso das Zusammenführen der formativen und finalen Evaluation (Wanner et al. 2018, S. 105). Die Co-Interpretation erfolgt hingegen gemeinsam mit Wissenschaftler\*innen und Praxispartner\*innen (ebd., S. 105). Hierzu sollten die Wissenschaftler\*innen jedoch einen geeigneten Rahmen schaffen bzw. gemeinsame Treffen organisieren (ebd., S. 105). Vgl. dazu Wanner et al (2018, S. 105): „The researchers’ role is to provide an overall evaluation concept, prepare evaluation meetings and connect formative and final evaluation“ (ebd., S. 105). In Rose et al. (2019, S. 9ff.) wird nochmals genauer auf Erläuterungen sowie Schlüsselerfahrungen und Ausprägungen aller Prozessschritte eingegangen. Aufgrund des Schwerpunkts dieser Arbeit auf Evaluation, soll der Fokus im Folgenden jedoch auf die Spezifika bzw. Erfahrungen im Rahmen der Co-Evaluation, wie sie hier konzipiert wird, eingegangen werden. Wie auch in Abb. 2 **Fehler! Verweisquelle konnte nicht gefunden werden.** zu sehen, besteht diese aus der Erfassung konkreter Ergebnisse, der Co-Interpretation und dem Transfer in Wissenschaft und Praxis (Wanner und Stelzer 2019, S. 3).



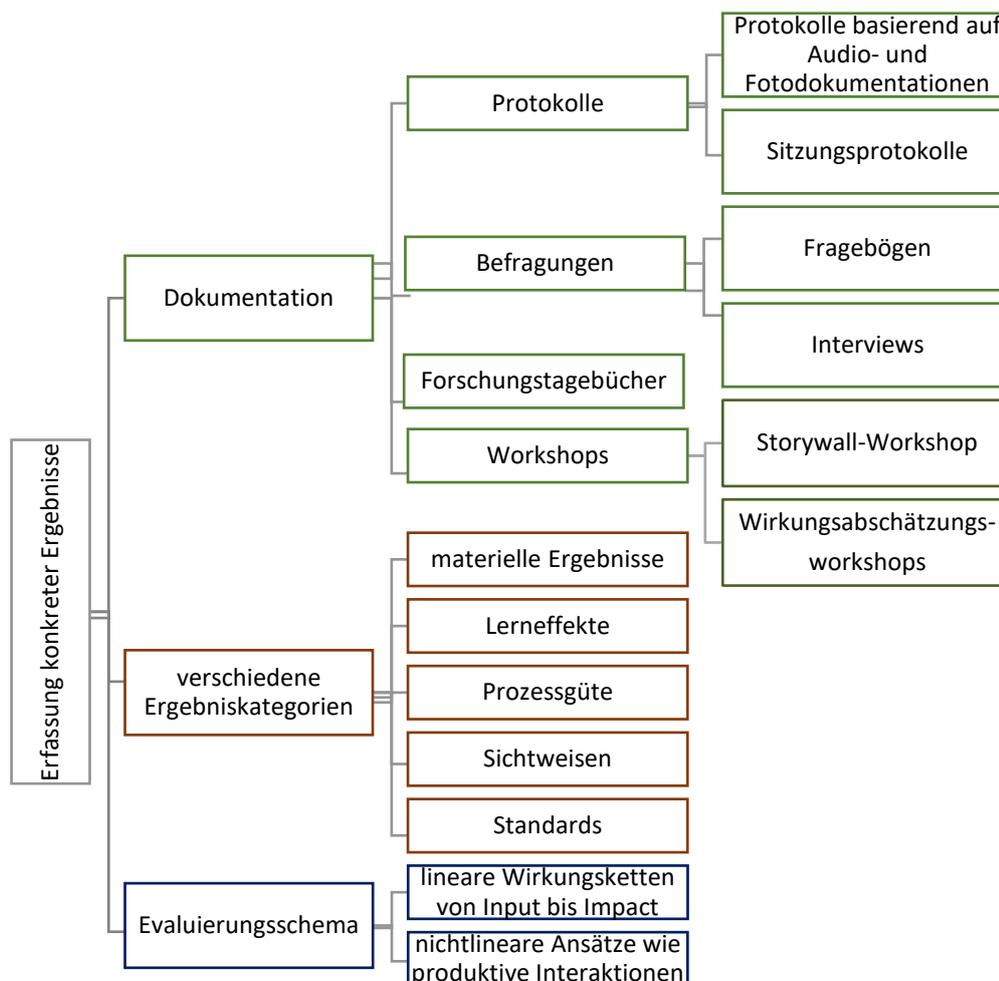
**Abb. 2: Das zirkuläre Reallabormodell.** Zu sehen sind die Prozesse des Co-Designs, der Co-Produktion und der Co-Evaluation und wie sie zyklisch miteinander verbunden sind. Die Vorgänge sind hier chronologisch von oben nach unten zu lesen und umfassen die Bereiche Wissenschaft und Praxis, welche beide durch transdisziplinäre Zusammenarbeit miteinander verbunden sind. Insgesamt ist das Reallabor in einen normativen Rahmen der nachhaltigen Entwicklung eingebettet (eigene Darstellung nach Wanner und Stelzer 2019, S.3, Original von Wanner et al. 2018, S. 102).

### 5.1.1.1 Erfassung konkreter Ergebnisse

Die Erfassung konkreter Ergebnisse umfasst drei Teilschritte, die Festlegung der Dokumentationsmethode(n), der Ergebniskategorien und des Evaluierungsschemas.

#### Dokumentation

Für die Dokumentation wurden in den von Rose et al. (2019) betrachteten Reallaboren (WTW, KInChem, WM) die in Abb. 3 dargestellten Methoden der Protokollierungen, Befragungen, Führung von Forschungstagebüchern sowie Workshops gewählt (Rose et al. 2019, S. 11). Diese Methoden können nochmals weiter aufgefächert werden in Protokolle basierend auf Audio- und Fotodokumentationen, Sitzungsprotokolle, Fragebögen, Interviews, Storywall-Workshops sowie Wirkungsabschätzungsworkshops (ebd., S. 11). Die Wahl der Methoden schien teilweise auf den jeweiligen Zeitraum und Zeitpunkt der Dokumentation angepasst worden zu sein. Im WM-Reallabor fanden beispielsweise vor der eigentlichen Durchführung Befragungen der Bewohner\*innen statt. Während der Durchführungsphase selbst wurden Memos angefertigt und zum Abschluss Nutzungsbefragungen sowie Prozessbewertungen mit den Praxispartner\*innen und Wissenschaftler\*innen durchgeführt (ebd., S. 11).



**Abb. 3: Erfassung konkreter Ergebnisse im Kontext der Co-Evaluation.** Die hier aufgeführten Aspekte spielen eine Rolle, wenn Daten für die Evaluation gesammelt werden sollen, wobei nach verschiedenen Arten der Dokumentation, Ergebniskategorien und Evaluierungsschemata gegliedert werden kann (eigene Darstellung nach Rose et al. 2019, S. 11).

### Verschiedene Ergebniskategorien

Nachdem hier bereits von „Ergebniskategorien“ gesprochen wird, erfolgt die Auswahl der Kriterien scheinbar auf Basis einer Produkt- bzw. Ergebnisevaluation, wie sie in Tab. 2 bereits kurz vorgestellt wurde (siehe Kapitel 4.3). Konkret werden hier materielle Ergebnisse, Lerneffekte, Prozessgüte, Sichtweisen und Standards als mögliche Kategorien angesprochen (Rose et al. 2019, S. 11). Wie in Kapitel 4.3 bereits erwähnt, ist es bei der Auswahl der Kriterien wichtig, dass sie sich empirisch erfassen lassen. Dementsprechend wirken sich die Ergebniskategorien wiederum auf die Auswahl der Dokumentationsmethoden aus, da entschieden werden muss, wie sich diese empirisch erfassen lassen.

### Evaluierungsschemata

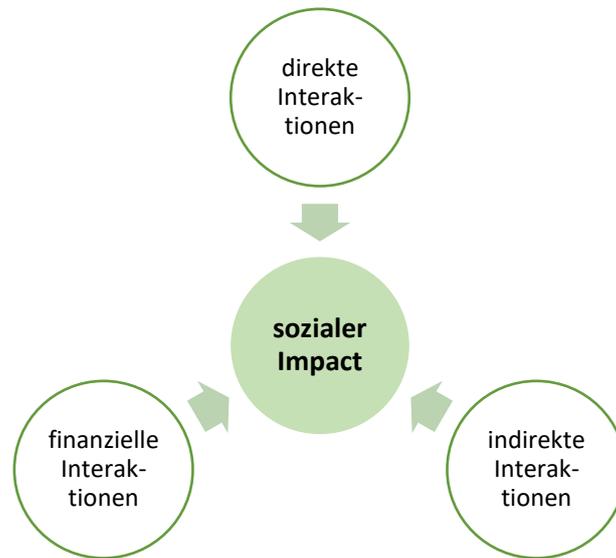
Das Evaluierungsschema bestimmt die Logik, wie die dokumentierten Daten miteinander in Verbindung gesetzt werden. Ein weitverbreitetes Model ist das der linearen Wirkungskette (Wanner et al. 2018, S. 105). Wie in Abb. 4 dargestellt besteht dieses aus fünf Sequenzen, nämlich den Inputs, den Prozessen bzw. Aktivitäten, den Outputs, den Outcomes sowie dem sozialen Impact (ebd., S. 105).



**Abb. 4: Modell der linearen Wirkungskette.** Dargestellt sind die fünf charakterisierenden, aufeinander folgenden Sequenzen des linearen Evaluierungsschemas und eine kurze Erklärung, was diese beinhalten (eigene Darstellung nach Wanner et al. 2018, S. 105).

Wanner et al. (2018, S. 105) und Rose et al. (2019, S. 11) weisen darauf hin, dass es neben dem vorgestellten linearen Evaluationsschema auch non-lineare Alternativen gibt. So verweisen die Autor\*innen beispielsweise auf den prozessorientierten Ansatz der „produktiven Interaktionen“ (ebd., S. 11; Wanner et al. 2018, S. 11). Als produktive Interaktionen werden nach Spaapen und van Drooge (2011, S. 212) ein Austausch zwischen Forscher\*innen und anderen Stakeholdern verstanden, in welchem Wissen produziert wird, das wissenschaftlich robust und sozial relevant ist (ebd., S. 212). Dabei können drei Arten produktiver Interaktionen unterschieden werden: direkte Interaktionen, indirekte Interaktionen und finanzielle Interaktionen (ebd., S. 213). Direkte Interaktionen, sind persönliche Wechselbeziehungen zwischen zwei Personen, zum Beispiel in Form von Face-to-Face-Kommunikation, Telefongespräche, E-Mail oder Videokonferenzen (ebd., S. 213). Indirekte Interaktionen entstehen, wenn über materielle Träger (carrier) kommuniziert wird, z.B. über Texte, Ausstellungen, Modelle oder Filme (ebd., S. 213). Finanzielle Interaktionen entstehen, wenn potentielle Stakeholder in einen ökonomischen Austausch mit Forscher\*innen treten, beispielsweise über einen Forschungsvertrag, einen finanziellen oder einen Sachbeitrag (ebd., S. 213). Als produktiv wird die Interaktion dann angesehen, wenn sie bei Stakeholdern dazu führt, dass sie Bemühungen anstellen, die wissenschaftlichen und praktischen Erkenntnisse oder Erfahrungen nutzen zu wollen (ebd., S. 212). Die Logik dahinter ist, dass produktive Interaktionen ein Voraussetzung sind für sozialen Impact, denn „in order to have impact you’ve got to have contact – direct, indirect and/or financial“ (ebd., S. 213). Durch den Fokus auf die produkti-

ven Interaktionen können verschiedene Varianten an Beteiligung differenziert werden (Spaapen und van Drooge 2011, S. 214). All diese Beteiligungsformen zusammengenommen sind wiederum Teil des Prozesses, welcher zu sozialem Impact führen kann (ebd., S. 212).



**Abb. 5: Formen produktiver Interaktionen.** Es existieren drei verschiedene Formen produktiver Interaktionen, welche gemeinsam zu sozialem Impact beitragen können (eigene Darstellung nach Spaapen und van Drooge 2011, S. 213).

### 5.1.1.2 Co-Interpretation

Während der Rahmen für die Evaluation, wie im vorherigen Kapitel angesprochen, von den Forscher\*innen organisiert wird, sollte die Auswertung in einem partizipativen Verfahren mit den Praxisakteuren stattfinden (Wanner et al. 2018, S. 105). Hierfür sollten nach Wanner et al. (2018, S. 105) verfügbare Leitlinien zu kooperativer Reflexion genutzt werden.

Ansätze zu partizipativen Auswertungsverfahren, wie es die Co-Interpretation bzw. kooperative Reflexion sind, finden sich zum Beispiel bei Unger (2014). Dabei beschreibt sie den Reflexions- und Auswertungsprozess als co-konstruktiven Prozess, dessen Aufgabe darin besteht „herauszufinden, wie und wie viel Partizipation jeweils in der Analyse ermöglicht werden kann, wie also die beteiligten Wissenschaftler/innen, Community- und Praxispartner zusammenarbeiten können“ (ebd., S. 61). Im Zuge dessen müssen nicht alle Schritte der Datenaufbereitung und -auswertung von allen Beteiligten persönlich durchgeführt werden (ebd., S. 63). Stattdessen können arbeitsteilige Vorgehen gewählt werden (ebd., S. 63). Dabei ist es jedoch wichtig transparent und nachvollziehbar zu kommunizieren, wie das Verfahren gestaltet ist (ebd., S. 63). Die von Arbeitsgruppen oder Einzelpersonen erarbeiteten Ergebnisse sollten zudem in einen gruppenübergreifende Diskurs und eine kommunikative Validierung münden (ebd., S. 63). Hierbei ist zu beachten, dass im Zuge der Arbeitsteilung nicht nur die Wissenschaftler\*innen die Auswertung vornehmen und im Anschluss zur Diskussion in der Gesamtgruppe freigeben, sondern die Auswertung auch aktiv durch weitere Projektteilnehmer\*innen vorgenommen wird (ebd., S. 63). Ansonsten besteht das Risiko einer zu einseitigen Auswertung aus akademischer Perspektive (ebd., S. 63). Im Zuge der partizipativen Auswertung soll aber genau die Diversität der verschiedenen Perspektiven der Projektbeteiligten Eingang finden.

Das Prinzip der Co-Interpretation erinnert in Hinblick auf eine gemeinsame Beurteilung an das Verfahren der diskursiven Evaluation von Bergmann et al. (2005). Die Autor\*innen haben einen Leitfaden für die Evaluation transdisziplinärer Projekte konzipiert und gehen ebenfalls auf einen Bewertungsprozess mit verschiedenen Beteiligten ein. Abhängig von der Projektgröße und -relevanz muss zunächst entschieden werden, ob alle oder nur ein Teil der Projektbeteiligten am Evaluationsprozess beteiligt sein sollen (ebd., S. 13). So können z.B. auch nur Vertreter\*innen der beteiligten Institutionen partizipieren oder aber auch noch externe Personen beteiligt werden (ebd., S. 13). Im Evaluationsverfahren selbst werden die Evaluationsfragen dann zunächst von den jeweiligen Einzelpersonen beantwortet. Im Anschluss daran sollen aber ebendiese individuellen Antworten mit den anderen Beteiligten ausgetauscht und besprochen werden (ebd., S. 13). Dabei „wird von Evaluierenden und Evaluierten gemeinsam ein analytischer Diskurs über die Hintergründe von Erfolgen und Misserfolgen des betrachteten Forschungsvorhabens geführt“ (ebd., S. 12). Das Ziel ist so zu einem gemeinsamen Ergebnis zu kommen (ebd., S. 13). Dieses Verfahren ist demnach im Partizipationsgrad sehr variabel, je nach beteiligten Personen. Durch den diskursiven Ansatz bietet sich insgesamt jedoch die Möglichkeit mittels der Beteiligung verschiedener Projektbeteiligter eine perspektivische Breite zu erfassen in Hinblick auf den gemeinsamen Auswertungsprozess.

In den von Rose et al. (2019, S. 14) betrachteten Reallaboren wurde die Co-Interpretation stets in Wissenschafts-Praxis-Workshops durchgeführt. Diese waren teilweise leitfadensbasiert (ebd., S. 14), wobei zur Art des Leitfadens keine genaueren Angaben getätigt wurden.

### 5.1.1.3 Transfer in Wissenschaft und Praxis

Im letzten Schritt der Co-Evaluation erfolgt der Transfer der Erkenntnisse in Wissenschaft und Praxis. Dafür werden die Ergebnisse in einen breiteren Kontext eingebettet und Schlussfolgerungen gezogen (Wanner et al. 2018, S. 106). Für den Transfer in die Wissenschaft bedeutet dies die Interpretation der Ergebnisse in Hinblick auf die Forschungsfrage, wissenschaftliche Theorien, Konzepte, Modelle und Systemanalysen (ebd., S. 106). Dadurch können die Ergebnisse in den Stand der Forschung sowie existierende theoretische und methodische Debatten integriert werden (ebd., S. 106). Die Verbreitung selbst erfolgt beispielsweise durch Beiträge auf Konferenzen oder in Publikationen (Rose et al. 2019, S. 12). Für den Transfer in die Praxis müssen zunächst die Lessons Learned zusammengefasst werden, wofür Erfolgs- und Misserfolgsmuster herausgearbeitet und abstrahiert werden sollten (Wanner et al. 2018, S. 105). Die daraus gewonnen Erkenntnisse können anschließend in verschiedenen Formen verbreitet werden (siehe Abb. 6).



**Abb. 6: Transfer in Wissenschaft und Praxis.** Dargestellt ist das Ablaufschema des letzten Teilschritts der Co-Evaluation, in welchem die gesammelten und co-interpretierten Erkenntnisse wieder in Wissenschaft und Praxis einfließen sollen (eigene Darstellung nach Wanner et al. (2018, S. 105f.) und Rose et al. (2019, S. 12)).

### 5.1.2 Begleitende externe Evaluation durch ein Evaluationsgremium

Während bei der projektinternen Co-Evaluation Evaluationskonzept und -rahmen von den Wissenschaftler\*innen entwickelt werden soll (Wanner et al. 2018, S. 105), sprechen sich Parodi et al. (2018, S. 5) für ein externes Evaluationsgremium aus, wie es bei den BaWü-Labs eingesetzt wurde. Dieses sollte folgende Eigenschaften aufweisen (ebd., S. 5):

- Das Evaluationsgremium sollte multidisziplinär zusammengesetzt sein.
- Die Gutachter\*innen sollten Erfahrungen mit transdisziplinärer und transformativer Forschung ausweisen, um den reallaborspezifischen Eigenheiten angemessen Rechnung tragen zu können.
- Die Begleitung sollte in gleicher Besetzung über die gesamte Laufzeit, von der Ausschreibungskonzeption ausgehend, erfolgen, um für Kontinuität zu sorgen.
- Den verschiedenen Aufgaben kann sich das Evaluationsgremium als Ganzes oder in Ausschüssen annehmen.

Zudem wird es als hilfreich erachtet, wenn die Gutachter\*innen selbst Erfahrungen darin hatten Reallabore umzusetzen sowie Personen aus der Organisationstheorie und -beratung einzusetzen, um auf die komplexen und dynamischen Strukturierungsprozesse in Reallaboren eingehen zu können (ebd., S. 5).

Doch auch wenn die Evaluation durch ein externes Evaluationsgremium durchgeführt wird, soll das Evaluationsverfahren „als kooperativer Lernprozess für Evaluierende [...] und Evaluierte gestaltet werden“ (ebd., S. 8). Dies soll erreicht werden, indem gemeinsam mit den Reallaboren ein Kriterienkatalog entwickelt wird, der den spezifischen Eigenheiten der Reallabore gerecht werden kann (Parodi et al. 2018, S. 8). Dieser dient anschließend als Grund-

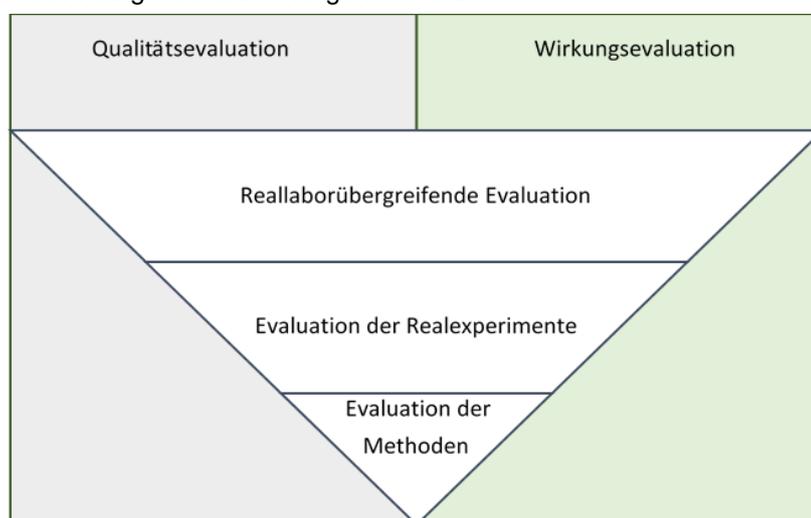
lage für die Evaluation (ebd., S. 8). Die Kriterien sollten des Weiteren wissenschaftliche Belange, das erkennbare Wirkungspotential (impact), Diffusionsaktivitäten, die Übertragbarkeit von Erkenntnissen sowie initiierte Lernprozesse einschließen (ebd., S. 8). Den jeweiligen Kriterien sollten ferner eindeutige Indikatoren zugeordnet werden (ebd., S. 9). Es wird empfohlen, transparent hinsichtlich der verwendeten Kriterien, Maßstäbe und Gewichtungen zu sein und deren jeweilige Verwendung zu erläutern. Dies ist aus Sicht von Parodi et al. (2018, S. 9) insbesondere deshalb wichtig, da sowohl Reallabore als Forschungsformat als auch die angemessene Evaluation ebenjener noch kontrovers diskutiert werden. Aufgrund der vorherrschenden Debatten und der fehlenden langjährigen Evaluationspraxis von Reallaboren wird dementsprechend dafür plädiert, eine angemessene Kritikoffenheit gegenüber den verwendeten Evaluationskriterien zu pflegen (ebd., S. 9).

Defila und Di Giulio (2018d, S. 102) betonen des Weiteren die Trennung von Begleitforschung und Evaluation, welche im Auftrag der Fördergebenden geschieht. Die Autor\*innen begründen dies dadurch, dass im Rahmen der Begleitforschung interne, geheim zu haltende Informationen ausgetauscht werden, die keinesfalls an Dritte, wie die Fördergebenden, weitergereicht werden dürfen (ebd., S. 102). Auch sollte eine Weiterleitung von internen Informationen über die Fördergebenden an die zu begleitenden Projekte vermieden werden (ebd., S. 102). Um die sensiblen Informationen zu schützen und Rollenkonflikte zu vermeiden, sollten Evaluation und Begleitforschung folgerichtig getrennt werden: „Accompanying research and evaluation should be treated as mutually exclusive activities“ (ebd., S. 102). Nichtsdestoweniger ist es möglich, dass die Begleitforschung die Verantwortung dafür übernimmt Kriterien und Partizipationsprozesse für die Evaluation zu entwickeln, (ebd., S. 102). Die tatsächliche Durchführung der Evaluation und die damit verbundene Bewertung der Reallabore bleibt dennoch strikt getrennt von der Begleitforschung (ebd., S. 102). In diesem Rahmen wurde auch die Entwicklung der Evaluationskriterien in den BaWü-Labs durchgeführt. So haben dort die Forscher\*innen der Basler Begleitforschungsgruppe „Reallabore vernetzen, verstehen, verstetigen“ die Aufgabe übernommen die Evaluationskriterien und die dazugehörigen partizipativen Prozesse zu entwickeln (ebd., S. 102). Dafür haben sie einen partizipativen Prozess entwickelt, initiiert und in diesem gemeinsam mit den Gutachter\*innen und den Mitgliedern der Reallabore einen Kriterienkatalog erarbeitet (ebd., S. 102). Damit hat die Begleitforschungsgruppe zwar den Grundstock für die Evaluation gelegt, ebendiese jedoch nicht selbst durchgeführt.

Anders als bei der Evaluationskonzeption von Wanner et al. (2018) wird Evaluation in den BaWü-Labs vor allem als Evaluation durch ein externes Evaluationsgremium kommuniziert. Dennoch spielen auch hier partizipative Prozesse eine Rolle, wie im vorherigen Abschnitt veranschaulicht wurde. Es handelt sich zwar nicht um eine Co-Interpretation der Evaluationsergebnisse, wie in der Konzeption der Co-Evaluation, sondern um eine partizipative Entwicklung der Evaluationskriterien, wie Defila und Di Giulio (2018d) beschreiben. Diese ging jedoch nicht von den Mitgliedern der Reallabore selbst aus, sondern wurde von den Wissenschaftler\*innen der Begleitforschung initiiert. Die Durchführung der Evaluation selbst wird, anders als bei der Co-Evaluation, von einem externen Evaluationsgremium durchgeführt. Dementsprechend handelt es sich um zwei perspektivisch unterschiedliche Evaluationsansätze, wobei die Co-Evaluation eher intern organisiert ist und die Evaluation nach dem Modell der BaWü-Labs von externen Evaluator\*innen sowie den Wissenschaftler\*innen der Begleitforschung organisiert wurde. Dabei müssen die Konzeptionen sich nicht zwangsläufig ausschließen, sondern können gegebenenfalls auch komplementär wirken.

### 5.1.3 Verschiedene Maßstäbe der Evaluation

In den vorherigen Kapiteln wurden die zwei Konzeptionen „interne Co-Evaluation“ und „externe Evaluation durch ein Evaluationsgremium“ vorgestellt. In diesem Kapitel sollen weitere Konzepte oder Herangehensweisen vorgestellt werden, die für die Evaluation in Reallaboren eingesetzt werden können. Nachdem sich aus der Literatur nach unserer Auffassung kein klarer Konsens hinsichtlich der Evaluation in Reallaboren ableiten lässt bzw. Informationen dazu eher dispers vorliegen, wird in der Folge versucht eine erste Strukturierung vorzunehmen. Dabei wird kein Anspruch auf Vollständigkeit erhoben, es sollen vielmehr verschiedene Ansätze zusammengeführt und damit ein erster Überblick über diverse Evaluationsmöglichkeiten geschaffen werden. Für diesen Zweck wird nach dem Maßstab bzw. dem konkreten Gegenstand der Evaluation unterschieden. Denn beim Sichten der Literatur fiel auf, dass es verschiedene Ebenen der Evaluation in Reallaboren gibt, welche jedoch bislang nicht explizit herausgearbeitet wurden. Deshalb wurde, wie in Abb. 7 dargestellt, eine Gliederung nach reallaborübergreifender Evaluation, Evaluation der Realexperimente und Evaluation der angewendeten Methoden differenziert. Zudem wurde eine Unterscheidung zwischen Qualitäts- und Wirkungsevaluation vorgenommen.



**Abb. 7: Einteilung der Evaluationsgegenstände nach Maßstab und Evaluationsart.**

Dieses Schema dient dazu Evaluationskonzeptionen danach sortieren zu können, welcher Maßstab dem betrachteten Evaluationsgegenstand zugrunde liegt und danach, ob sie ihren Fokus auf die Qualitätssicherung des betrachteten Evaluationsgegenstandes oder auf dessen Wirkung legt (eigene Darstellung).

Als reallaborübergreifende Evaluation werden (Teil-)Evaluationen oder Evaluationskriterien angesehen, die sich auf das Reallabordesign, die übergeordnete Infrastruktur, übergreifendes Wissen o.Ä. beziehen. Hier wird beispielsweise das Reflexionsschema von Beecroft et al. (2018, S. 94f.) eingeordnet (siehe Anhang, Tab. 9). Dieses stellt eine Matrix aus Forschungszielen, Praxiszielen und Bildungszielen dar, welchen die Designprinzipien „Problem- und Themenangemessenheit herstellen“, „räumliche Angemessenheit gestalten“, „zeitliche Angemessenheit herstellen“, „angemessene Akteursrollen etablieren“ und „experimentell-reflexive Arbeitsweise fördern“ quergelegt sind (ebd., S. 94). Das Schema wurde zwar nicht direkt für den Evaluationsprozess erstellt, kann den Autor\*innen zufolge aber auch hierfür genutzt werden (ebd., S. 96). Dabei wird vorgeschlagen die leere Tabelle dazu zu nutzen, die verschiedenen Perspektiven der involvierten Akteure zu erheben. Auf Grundlage dessen kann das Design des Reallabors schrittweise abgestimmt werden (ebd., S. 96). Dabei würde

es sich also um einen formativen Evaluationsansatz handeln, der dazu dient, das Reallabor-design in Hinblick auf die gewünschten Forschungsziele zu optimieren.

Auf der nächsten dargestellten Maßstabsebene befindet sich die Evaluation der Realexperimente. Hierzu haben Luederitz et al. (2017) eine eigenständige Publikation veröffentlicht. Dabei stellen sie ein eigenständiges, prototypisches Evaluationsschema für die Evaluation von „sustainability transition experiments“ vor, welche hier synonym zu Realexperimenten verstanden werden. Luederitz et al. (2017, S. 62) verstehend diese als Evaluationsgegenstand wie folgt:

1. Realexperimente definieren eine Grundlinie (baseline) und das Ziel (goal) für ihre Evaluation,
2. erschaffen ein konkretes Set-up, um Interventionen zu verwalten,
3. vergleichen die Effekte der Interventionen mit der Grundlinie und dem Ziel,
4. vergleichen die Effekte mit den Nachhaltigkeitskriterien
5. und bieten belegbare Empfehlungen für Mainstream-Lösungen.

Das von ihnen entwickelte Evaluationsschema soll dazu dienen das Ausmaß zu bewerten, in welchem das Realexperiment die gewünschten Effekte generiert hat und wie dies erreicht wurde (ebd., S. 64). Gegliedert ist das Schema nach den vier Dimensionen „Outputs“, „Outcomes“, „Inputs“ und „Prozesse“. Luederitz et al. (2017, S. 64) haben dabei die Sequenz von der chronologischen Ablauflogik der Experimentdurchführung (Inputs – Prozesse – Outputs – Outcomes) in die Logik der Evaluationsdurchführung überführt. Das bedeutet, dass das primäre Interesse in den Outputs und Outcomes liegt, um davon ausgehend die Prozesse und Inputs rückwirkend bewerten zu können (ebd., S. 64). Die Sequenzen werden in der Praxis jedoch nicht als chronologisch, sondern als parallel und interdependent ablaufend angesehen (ebd., S. 64). Dies erfordert iterative Evaluationsprozesse innerhalb der vier Evaluationsdimensionen (ebd., S. 64). Um den Realexperimenten die Möglichkeit zu bieten, sich mittels der Evaluation entsprechend anzupassen, auszuweiten oder zu verändern, nennen Luederitz et al. (2017, S. 62) neben den iterativen Evaluationsprozessen des Weiteren die Notwendigkeit ex-ante, formative und ex-post Evaluationen durchzuführen. Die ex-ante Evaluation wird vor der Implementationsphase des Experiments durchgeführt, um das Design zu evaluieren (ebd., S. 62). Die formative Evaluation beschreibt die Evaluationsschleifen zur Anpassung und Verbesserung laufender Experimente (ebd., S. 62). Die abschließende ex-post Evaluation dient dazu, den Beitrag des Experiments zur Nachhaltigkeit nach dessen Durchführung zu bewerten (ebd., S. 62). Das hierfür entwickelte Evaluationsschema findet sich im Anhang in Tab. 10, wobei zusätzlich zu den Kriterien und Evaluationsfragen noch die im Fließtext genannten möglichen Indikatoren aufgenommen wurden. Diese sind den Autor\*innen zufolge jedoch noch nicht vollständig operationalisiert, sondern sollen als Platzhalter und Gedankenstütze dienen bis das Evaluationsschema konkret Anwendung findet (ebd., S. 63).

Die kleinste Maßstabsebene in der hier vorliegenden Systematik ist die Evaluation der Methoden. Auch hierfür kann das oben genannte Reflexionsschema von Beecroft et al. (2018) verwendet werden. So kann das Schema dafür genutzt werden eine Methode dahingehend zu prüfen, inwiefern sie adäquat für das jeweilige Reallabor ist oder ob ihr Potenzial vollständig ausgenutzt wird (ebd., S. 96). Des Weiteren haben Defila und Di Giulio (2018a) einen Katalog von Gütekriterien für Methoden der partizipativen Wissenserzeugung entwickelt. In diesem werden die Kriterien nicht konkret für den Zweck der Evaluation entwickelt, sondern

sollen, „bei der Wahl (inkl. Entwicklung und Kombination) und bei der Umsetzung von Methoden der partizipativen Wissenserzeugung sowie bei der Dokumentation und Reflexion des methodischen Vorgehens berücksichtigt werden“ (Defila und Di Giulio 2018a, S. 61). Weiterführend sagen sie jedoch auch, dass sich noch bewähren muss, in welchen Kontexten die Kriterien Anwendung finden können (ebd., S. 61). Nachdem die kontextuelle Einbettung noch nicht eindeutig ist, sollen sie hier auch in die Debatte der Evaluation von Reallaboren eingebracht werden. Dabei könnten die Gütekriterien dazu dienen, die verwendeten Methoden im Sinne einer Qualitätsevaluation zu evaluieren und damit zu bewerten, ob ihre Anwendung angemessen ist. Im Anhang (Tab. 11) finden sich die vorgestellten Gütekriterien und Erläuterungen in Tabellenform. Die von Defila und Di Giulio (2018a) aufgeführten Erläuterungen der Kriterien wurden in der vorliegenden Arbeit in Fragen umgewandelt (ebd., S. 55), mit dem Ziel, die verschiedenen hinter den Kriterien liegenden Aspekte zu dekonstruieren und so eine mögliche Indikatorenbildung zu erleichtern.

## 5.2 Untersuchung der BaWü-Labs

### 5.2.1 Methodik

Ziel dieser Arbeit ist es herauszuarbeiten, wie Reallabore bisher evaluiert wurden, um darauf aufbauend Ansatzpunkte für die Umsetzung im Projekt Innoredux herauszuarbeiten. Nachdem in den Kapiteln zuvor bereits bisherige Evaluationsansätze aus der Literatur besprochen wurden, soll nun darauf eingegangen werden, mit welcher Methodik die Evaluationspraxis in den BaWü-Labs genauer untersucht wurde. Die Ergebnisse daraus werden schließlich im Folgekapitel dargestellt.

Die BaWü-Labs wurden einerseits gewählt, da sie einen geographischen Bezug auf Länderebene zu Innoredux aufweisen, da all diese Reallabore in Baden-Württemberg angesiedelt sind. Andererseits bieten die Reallabore der beiden Förderlinien eine gute Forschungsgrundlage, da sie mit 14 Reallaboren einen thematisch diversen, aber doch überschaubaren und sehr konkreten Rahmen bieten. Wie in 5.1.2 beschrieben werden zudem alle Reallabore von einem externen Gutachtergremium evaluiert. Nachdem bei Parodi et al. (2018), Zimpelmann (2018) oder Defila und Di Giulio (2018d) vor allem programmübergreifend über Evaluationsaspekte der BaWü-Labs gesprochen wird, sollte hier die Evaluationspraxis der einzelnen Reallabore genauer in den Blick genommen werden. Damit sollen tiefere Einblicke und eine Übersicht über verschiedene Herangehensweisen geschaffen werden, um für spätere Evaluationsvorhaben eine Grundlage zu bieten.

Um dies zu erreichen wurde eine qualitative Literaturrecherche durchgeführt. Hierfür wurde zunächst in der Datenbank von google scholar nach relevanter Literatur gesucht, wobei sowohl BaWü-Lab spezifische Begriffe als auch allgemein Reallabor-bezeichnende Begriffe verwendet wurden. Konkret wurde nach den folgenden Begriffen gefiltert: ‚Reallabor Evaluation‘, ‚Reallabor evaluieren‘, ‚Evaluationskonzept Reallabor‘, ‚Evaluationskriterien Reallabor‘, ‚real-world laboratory evaluation‘, ‚real-world laboratory evaluating‘, ‚real-world laboratory evaluate‘, ‚real-world laboratory evaluation concept‘, ‚BaWü-Lab evaluation‘, ‚BaWü-Lab evaluieren‘, ‚BaWü-Lab Evaluationskonzept‘, ‚BaWü-Lab Evaluationskriterien‘. Zusätzlich wurden noch die Projektseiten der BaWü-Labs durchsucht. Dabei wurden sowohl Publikationen als auch graue Literatur beachtet. Die Datensammlung wurde von April bis Mai, im Jahr 2020 durchgeführt, wodurch nur Ergebnisse berücksichtigt werden konnten, die in diesem Zeitraum auf den Webseiten verfügbar waren.

Die gesammelten Daten wurden schließlich in eine Matrix eingepflegt, die sich an den von Stockmann und Meyer (2014, S. 75) formulierten Leitfragen „Was (welcher Gegenstand) wird wozu (zu welchem (Zweck) anhand welcher Kriterien von wem wie (mit welchen Methoden) evaluiert?“ orientiert. Dabei wurden die jeweiligen Leitfragen als Spaltenkategorien den Reallaboren zugeordnet. Zudem wurde die Kategorie „Wann wurde die Evaluation durchgeführt?“ aufgenommen, um für die spätere Praxisanwendungen auch Informationen über den geeigneten Durchführungszeitpunkt erarbeiten zu können. Die Informationen aus den Literaturquellen wurden hierbei zunächst als wörtliche Zitate in die jeweiligen Zeilen und Spalten übertragen und anschließend generalisiert und vereinfacht, um eine bessere Übersichtlichkeit zu gewinnen. Dieses Vorgehen wurde gewählt, da sich die Leitfragen als sehr übersichtliche und intuitive Kategorien umsetzen ließen, aber auch da diese wiederum der Ausgangspunkt für neue Evaluationen sind. Damit wird versucht die Nutzbarkeit der Ergebnisse für die Praxis zu erhöhen.

Durch die genaue Beschreibung des Suchvorgangs sowie einer Auflistung der gefundenen Literaturquellen (siehe Tab. 3) wurde versucht, eine größtmögliche Transparenz zu erreichen, die es ermöglichen soll gegebenenfalls übersehene Quellen oder Recherchemöglichkeiten identifizieren zu können. Denn durch die manuelle und einseitig durch eine Person durchgeführte Literaturrecherche kann nicht ausgeschlossen werden, dass trotz hoher Sorgfalt auch Informationen übersehen wurden, insbesondere, wenn diese in Literaturquellen implementiert sind, deren Titel und Beschreibung nicht zwingend auf Angaben zur Evaluation schließen lässt. Dennoch sollte mittels der zweigleisigen Recherche sowohl in der google scholar-Datenbank als auch der einzelnen Webseiten eine hohe Abdeckung der möglichen Angaben erreicht werden, um so eine zuverlässige Datenbasis für die Auswertung zu generieren. Tab. 3 zeigt, welche Informationen im Rahmen der Literaturrecherche zu welchem BaWü-Lab gefunden wurden. Die ausgegrauten Bereiche sind diejenigen Reallabore, zu welchen keine konkret zuordenbaren Informationen zur Evaluation gefunden wurden. Falls die Webseite nicht aufgerufen werden konnte, wurde dies ebenfalls vermerkt. Auf Grundlage dieser Quellen konnte schließlich die weitere Auswertung erfolgen, wie sie im folgenden Kapitel dargelegt wird.

**Tab. 3: Übersicht über die gefundenen Informationsquellen zu Evaluation in den BaWü-Labs.** Die vorliegende Tabelle zeigt eine Auflistung aller 14 BaWü-Labs, den Links zu den dazugehörigen Projektwebseiten, welche im Rahmen der Arbeit nach Informationen hinsichtlich der Evaluationspraxis durchsucht wurden, sowie Angaben zu den gefundenen Informationsquellen aus der Datenbank- und Webseitenrecherche. Zu den grau hinterlegten Reallaboren lagen zum Zeitpunkt der Recherche (Frühjahr 2020) keine genaueren Informationen vor (eigene Darstellung).

BaWü-Lab	Link zur Webseite	Gefundene Informationsquellen	Art der Information
WiNo	<a href="http://www.wissensdi-log-nordschwarz-wald.de/">http://www.wissensdi-log-nordschwarz-wald.de/</a>		
R 131: KIT findet Stadt	<a href="https://www.quartierzukunft.de/forschung/reallabor-131/">https://www.quartierzukunft.de/forschung/reallabor-131/</a>		
BaWü-Lab RNM	<a href="http://www.r-n-m.net/">http://www.r-n-m.net/</a>	Lazarova et al. (2018)	Ausführlichere Beschreibung

Dietenheim zieht an	<a href="https://www.uni-ulm.de/mawi/reallabor/">https://www.uni-ulm.de/mawi/reallabor/</a> <a href="http://dietenheim-zieht-an.de/">http://dietenheim-zieht-an.de/</a>	Geiger et al. (2017)	Konkrete Methodenbeschreibung
Urban Office	<a href="https://www.geog.uni-heidelberg.de/hca/reallabor.html">https://www.geog.uni-heidelberg.de/hca/reallabor.html</a>	Geographisches Institut (2018); Marquardt & West (2016); West (2018)	Kurzinformationen
EnSign	(Seite konnte nicht aufgerufen werden)		
Space Sharing	<a href="http://www.abk-stuttgart.de/forschung/forschungsprojekte/reallabor-space-sharing.html">http://www.abk-stuttgart.de/forschung/forschungsprojekte/reallabor-space-sharing.html</a>		
BUGA:log	(Seite konnte nicht aufgerufen werden)	Marsden et al. (2018) Marsden et al. (2019)	Kurzinformationen
Reallabor Schorndorf	<a href="https://www.reallabor-schorndorf.de/">https://www.reallabor-schorndorf.de/</a>	Klötzke, et al. (2018); Brandies et al. (2017); Schnieder & Gebhardt (2016); König et al. (2016)	Kurzinformationen und ausführliche Methodenbeschreibung
GO Karlsruhe	(Seite konnte nicht aufgerufen werden)	Häußler et al. (2019)	Kurzinformation
Energielabor Tübingen	(Seite konnte nicht aufgerufen werden)		
Stadt:quartiere 4.0	<a href="https://digitale-mitwirkung.de/">https://digitale-mitwirkung.de/</a>	Städtebau-Institut Stuttgart (2019)	Kurzinformation
STADT-RAUM-BILDUNG	<a href="https://stadt-raum-bildung.de/">https://stadt-raum-bildung.de/</a>		
Reallabor Asylsuchende in der Rhein-Neckar-Region	<a href="https://www.reallabor-asyl.de/reallabor-asyl/reallabor-asyl.html">https://www.reallabor-asyl.de/reallabor-asyl/reallabor-asyl.html</a>		

## 5.2.2 Ergebnisse

In diesem Kapitel sollen nun die gesammelten Erkenntnisse aus der durchgeführten Literaturrecherche dargelegt werden. In Tab. 4 wird dazu zunächst ein Überblick über die Kontexte der Reallabore gezeigt, zu welchen Informationen vorliegen, um diese besser verorten zu können. In Tab. 5 folgt die inhaltliche Darstellung der gesammelten Erkenntnisse über die Evaluation der Reallabore. Zuzüglich der Ergebnisse zu den BaWü-Labs wurden in Tab. 4 und Tab. 5 noch die Informationen zu den drei Reallaboren aus der Publikation von Rose et

al. (2019) eingefügt. Dabei handelt es sich zwar nicht um BaWü-Labs, aber nachdem im Einführungskapitel (siehe Co-Evaluation als Teil eines zirkulären Reallabormodells, s. Kap. 5.1.1) eingehender auf die verwendeten Methoden eingegangen wurde, wurden die Daten zum Zweck der Datenzusammenführung ebenfalls mit aufgenommen.

In Kapitel 5.1.2 wurde bereits die Evaluation durch das Evaluationsgremium beschrieben, welche ebenfalls die einzelnen BaWü-Labs sowie das Gesamtprogramm evaluiert haben (Zimpelmann 2018, S. 321). Wie die Auswertung der Literatur in der Matrix zeigt, erfolgten neben diesem externen Evaluationsansatz jedoch auch interne Evaluationen in den einzelnen BaWü-Labs. In den ausgewerteten Literaturquellen der BaWü-Labs wurde in sieben Fällen angegeben, dass im Reallabor durchgeführte Realexperimente bzw. konkrete Maßnahmen evaluiert wurden. Demnach zeigt sich Evaluation in den BaWü-Labs auf mehreren Maßstabsebenen - sowohl Reallabor- und Programmübergreifend durch das externe Evaluationsgremium als auch intern auf Ebene der einzelnen Maßnahmen. Es können allerdings keine Aussagen darüber getroffen werden, inwiefern die interne und externe Evaluation miteinander verbunden waren oder ob sie unabhängig voneinander durchgeführt wurden. Wie die grau markierten Bereiche zeigen (siehe Tab. 5), fehlen in sieben der insgesamt 14 untersuchten BaWü-Labs noch Angaben hinsichtlich der Evaluation in den Reallaboren. Es ist unklar, ob hier keine internen Evaluationen durchgeführt wurden, die Informationen diesbezüglich nicht öffentlich oder an anderer Stelle vorliegen oder ob hier lediglich Evaluationen durch das Evaluationsgremium erfolgt sind.

Die Schwerpunktsetzung in Bezug auf die Evaluationsfunktion zeigt sich unterschiedlich in den verschiedenen Reallaboren. Wie in Kapitel 4.2 beschrieben, kann eine Evaluation fünf verschiedene Funktionen innehaben, nämlich die Erkenntnisfunktion, Lern- und Dialogfunktion, Optimierungsfunktion, Entscheidungsfunktion oder die Legitimationsfunktion (Döring und Bortz 2016, S. 987). Aus der untersuchten Literatur lässt sich vorwiegend eine Erkenntnis- und/oder Optimierungsfunktion ableiten. So zeigt sich beispielsweise in den Reallaboren „Stadt:quartiere 4.0“, „Urban Office“, „Dietenheim zieht an“ und im „BaWü-Lab RNM“, dass Erkenntnisse darüber gewonnen werden sollen, inwiefern die jeweilige Maßnahme wirkt (Marquardt und West 2016, S. 29) oder genutzt wird und ob sie dabei erfolgreich (Geiger et al. 2017, S. 147; Lazarova et al. 2018, S. 178) bzw. geeignet ist (Städtebau-Institut Stuttgart 2019, S. 27). Die Optimierungsfunktion zeigt sich am besten im „Reallabor Schorndorf“, in welchem die Weiterentwicklung des Quartierbussystems bezweckt werden soll (König et al. 2016, S. 5). Aber auch im „Urban Office“ sollen Optimierungsmöglichkeiten durch die Evaluation geschaffen werden (Geographisches Institut Heidelberg 2018). Im Reallabor „Dietenheim zieht an“ zeigt sich des Weiteren auch eine Entscheidungsfunktion, da dort mittels der Evaluation Rückschlüsse gezogen werden sollen, was bei nachfolgenden Reallaboren gegebenenfalls in der Konzeption verändert werden soll (Geiger et al. 2017, S. 148). Um die intendierte, primäre Funktion der Evaluation zu ergründen, sollten jedoch noch weitere Nachforschungen betrieben werden. Auch in denjenigen Reallaboren, die entweder keine konkret zuordenbaren Aussagen diesbezüglich getroffen haben oder in welchen keine Informationen über mögliche Evaluationstätigkeiten vorliegen. Dahingehend könnte auch interessant sein, inwiefern die Evaluation intrinsisch durch das Projektteam motiviert war oder ob sie durch Fördergebende auferlegt wurde. Insgesamt lässt sich aus der Auswertung der hier vorliegenden Daten ableiten, dass Erkenntnis- und Optimierungsfunktionen eine Rolle gespielt haben in der Konzeption der Evaluation.

In zwei Reallaboren wurde angegeben, dass Erfolgskriterien verwendet wurden, konkret in „Dietenheim zieht an“ (Geiger et al. 2017, S. 147) und im „BaWü-Lab RNM“ (Lazarova et al.

2018, S. 177f.). In den anderen Reallaboren wurden keine oder keine aussagekräftigen Informationen über konkret verwendete Kriterien gefunden. Im „BaWü-Lab RNM“ wurden die Kriterien in einem Transformationsworkshop festgelegt (ebd., S. 17). Bei „Dietenheim zieht an“ wurde der Kriterienkatalog durch das Reallaborteam festgelegt, nachdem die involvierten Stakeholder ihre Zielvorstellungen geäußert hatten (Geiger et al. 2017, S. 147). In diesen beiden Vorgehen zeigen sich partizipative Vorgehen der Kriterienermittlung, ähnlich wie sie auch durch die Begleitforschung für die externe Evaluation vorgenommen wurden (Defila und Di Giulio 2018d, S. 102).

Insbesondere mit Blick auf die evaluierenden Personen wird deutlich, dass es in den BaWü-Labs noch weitere Ebenen, neben der externen Evaluation in den BaWü-Labs gibt. So wird in „Dietenheim zieht an“ bspw. durch das Forschungsteam evaluiert (Geiger et al. 2017, S. 148) und im „Urban Office“ erfolgt die Evaluation durch Studierende (Marquardt und West 2016). Damit entsprechen die internen Evaluationsansätze nicht dem Co-Evaluationsansatz von Wanner et al. (2018), da hier den Angaben zufolge nicht co-interpretiert, sondern vorwiegend aus akademischer Perspektive bewertet wird.

Hinsichtlich des Zeitpunkts der Evaluation wurde in drei Fällen erwähnt, dass es sich um wiederholte bzw. iterative Evaluationen handelt (König et al. 2016, S. 5; Marquardt und West 2016, S. 29; Marsden et al. 2019, S. 29). Auch wurden teilweise im Laufe des Durchführungsprozesses Anpassungen des Evaluationskonzepts vorgenommen (Lazarova et al. 2018, S. 178).

In Hinblick auf die verwendeten Methoden wird deutlich, dass vor allem qualitative Methoden genannt werden. In Tab. 6 werden diese zur besseren Übersicht in partizipative, beobachtende, befragende und sonstige Methoden sortiert, wobei hier auch die Reallabore von Rose et al. (2019) aufgenommen wurden. Durch die Sortierung wird deutlich, dass bei der Wahl der Evaluationsmethoden unterschiedliche Arten und Grade der Partizipation von Evaluierenden und Evaluierenden vorhanden sind.

**Tab. 4: Kontext der Reallabore, zu welchen Informationen über die Evaluation vorliegen.** Diese Matrix enthält Informationen über kontextuelle Rahmenbedingungen der betrachteten Reallabore, um eine grobe Einordnung ebenjener zu ermöglichen. Hierin sind nur die BaWü-Labs aufgelistet, zu welchen Informationen über ihre Evaluation vorlagen (grün). Zum Zweck der Datenzusammenführung wurden zusätzlich die Reallabore aus der Publikation von Rose et al. (2019) eingebunden (blau), da dort ebenfalls wichtige Erkenntnisse hinlänglich der Evaluation gefunden wurden. Zusätzlich ist das Projekt Innoredux aufgeführt (gelb) (eigene Darstellung).

Reallabor	Kontext			
	Ort	Anwendungsfeld	Übergeordnete Zielsetzung	Quelle
BAWü-Lab RNM	Stuttgart	Mobilität	Förderung einer nachhaltigen Mobilitätskultur, Schaffen neuer Lebens- und Aufenthaltsqualitäten in Stadt und Region	(Defila und Di Giulio 2018c, S. 379)
Dietenheim zieht an	Dietenheim	Textilindustrie	Nachhaltigkeitsorientierte Revitalisierung der Textilstadt	(Webseite Dietenheim zieht an)

Urban Office	Heidelberg	Stadt-entwicklung	Generieren von Wissen zu urbanen Transformationsprozessen im Kontext der Stadtentwicklung in der Wissensgesellschaft	(Geographisches Institut Heidelberg 2018)
BUGA:log	Heilbronn	Mobilität	Entwicklung eines automatisierten Gütertransportsystems	(Marsden et al. 2018, S. 2)
Reallabor Schorndorf	Schorndorf	Mobilität	Entwicklung eines flexiblen, bedarfsgerechten und digital gestütztem Bussystem	(Defila und Di Giulio 2019a, S. 352)
GO Karlsruhe	Karlsruhe	Mobilität	Förderung des Fußverkehrs	(Häußler et al. 2019, S. 396)
Reallabor Stadt:quartiere 4.0	Stuttgart und Herrenberg	Stadt-entwicklung	Untersuchung von Auswirkungen der Digitalisierung auf urbane Lebensräume und Entwicklung digitaler Instrumente für Planungs- und Entscheidungsprozesse	(Städtebau-Institut Stuttgart 2019, S. 1)
WTW	Wuppertal	Stadt- und Wirtschaftsentwicklung	Förderung eines ressourcenleichten Wohlstands auf Basis ökonomischer, ökologischer und sozialer Wohlstandsaspekte	(Webseite WTW o.J.)
WM	Pirmasens, Werder an der Havel, Berlin	Wohnen	Untersuchung und Erprobung gemeinschaftlicher Mobilitäts- und wohnungsbezogener Dienstleistungen	(Schönau et al. 2018)
KInChem	Keine Information vorhanden	Chemikalien	Erprobung von Informationsinstrumenten zur Gewinnung von Gestaltungsoptionen im Rahmen der EU-Chemikalien-Verordnung REACH	(NaWiKo o.J.)
Innoredux	Heidelberg	Verpackungen	Entwicklung und Erprobung von nachhaltigen Verpackungslösungen	(Webseite Innoredux o.J.)

**Tab. 5: Evaluationspraxis der betrachteten Reallabore.** Mittels dieser Matrix wurde die Auswertung der Literaturrecherche vorgenommen, wobei die gesammelten Informationen über die einzelnen Reallabore den Leitfragen nach dem Evaluationsgegenstand, -zweck, -kriterien, den verantwortlichen Evaluator\*innen, verwendeten Methoden und dem Zeitpunkt der Evaluation nach Stockmann und Meyer (2014, S. 75) zugeordnet wurden. Die gefundenen Informationsquellen können so eindeutig einer Leitfrage und einem Reallabor zugeordnet werden (eigene Darstellung).

	Was (welcher Gegenstand) wird evaluiert?	Wozu (zu welchem Zweck) wird evaluiert?	Anhand welcher Kriterien wird evaluiert?	Wer evaluiert?	Wie (mit welchen Methoden) wird evaluiert?	Wann wird evaluiert?
<b>WiNo</b>						
<b>R 131: KIT findet Stadt</b>						

<b>BaWü-Lab RNM</b>	Parklets (Lazarova et al. 2018)	Evaluation der Nutzung der Parklets; Untersuchung der Gründe dafür oder dagegen; Erfassung der öffentlichen Meinung (ebd., S. 26); Ziele auf Erfolg prüfen (ebd., S. 178)	Es gibt 4 Ziele mit insgesamt 12 Erfolgskriterien (ebd., S. 177f.)	Evaluation der Parklets durch Studierende (ebd., S. 195)	<p>Zweiteiliger <b>Transformationsworkshop, zur Festlegung der Ziele</b> der Experimente, <b>Evaluationskriterien</b> und Datenerhebungsmethodik (ebd., S. 17)</p> <p><b>Raum und Forschungstagebuch:</b> „Zählen von Parklet Nutzer*innen, Aufnahme von Bewegungsabläufen in der Umgebung der Parklets, Beschreibung des Umfeldes, der Ereignisse und der Aktivitäten der Nutzer*innen, Aufzeichnung des Wetters und Beobachtungszeiten sowie das Feedback von Paten*innen und Anwohnerinnen durch Beobachtungen und Gespräche vor Ort“ (ebd., S. 178)</p> <p>Zusätzliche <b>Beobachtungen</b>, da Tagebücher nicht alle gleich gut gepflegt wurden, (ebd., S. 178)</p> <p>zusätzliche <b>Leitfadengestützte Interviews</b> mit den Pat*innen, da Tagebücher nicht alle gleich gut gepflegt wurden (ebd., S. 178)</p> <p><b>Inhaltsanalyse</b> der Presseartikel, Facebookposts, Feedback-E-Mails etc. (quantitative Aussagen zur Aritkelanzahl und qualitative Aussagen in Bezug auf Begründungsmuster der Befürworter*innen und Gegner*innen) (ebd., S. 178)</p>	Transformationsworkshop (06/2016) ein Jahr nach Reallaborbeginn (03/2015) (ebd., S. 18f.); Anpassung des Evaluationskonzepts während des Prozesses (ebd., S. 178)
-------------------------	---------------------------------------	---	--	--	---	---

<b>Dietenheim zieht an</b>	Im Reallabor entwickelte Maßnahmen, die sich an die Einwohner(inne) der betroffenen Gemeinde richten (Geiger et al. 2017, S. 147)	Messen des (Teil-) Erfolgs der durchgeführten Maßnahmen (ebd., S. 147)  Hinweise darauf, was bei nachfolgenden Reallaboren mit ähnlichem Setting verändert werden müsste, falls keine Veränderungen sichtbar sind (durch Prä-Post-Design zu realisieren) (ebd., S. 148);  Qualitätssicherung (ebd., S. 148)	Erfolgskriterien-Katalog (ebd., S. 147)	Ausloten der Zielvorstellungen von Stakeholdern (ebd., S. 147);  Bestimmung des Erfolgskriterienkatalogs durch das Reallaborteam (ebd., S. 147);  Evaluation durch das Forschungsteam (ebd., S. 148)	Erhebungen bei den Veranstaltungen für die Erfassung des direkten Erfolgs (ebd., S. 148) Erfassung qualitativer Daten im Nähcafé (ebd., S. 148) <b>Beobachtungen</b> (zur Evaluation des Prozesses) (ebd., S. 148) <b>Halbstrukturierte Interviews</b> mit Teilnehmer*innen und Workshopleiter*innen (zur Evaluation des Prozesses) (ebd., S. 148) Austeilen von <b>Feedbackkarten</b> vor und nach jedem Workshop (zur Evaluation des Prozesses) (ebd., S. 148) <b>Prä-Post-Design</b> , um sich längerfristigeren Wirkungen anzunähern (Vergleich diverser Variablen vor und nach einer Maßnahmenintervention) mittels <b>Fragebogen</b> (Variablen beziehen sich auf die Erfolgskriterien und beinhalten einen allgemeinen Teil zur Zufriedenheit mit dem Heimatort und den Entwicklungswünschen; Befragung erfolgte im Abstand von 2 Jahren) (ebd., S. 148)
<b>Urban Office</b>	Bei „Wissen to go“-Format: Die mobilen Wissensorte (Musikwagen, mobiler Wissensspiel-	Wirkung der mobilen Wissensorte (Marquardt und West 2016, S. 29);  Zwischenevaluatio- nen: Gewinnen von Erkenntnissen über		Evaluation durch Studierende	Einführen von Beobachtungsebene neben der Handlungsebene: Evaluation der Interaktionen mit den Praxispartner*innen, d.h. den Akteuren bei der Konzeptualisierung und beim Umbauen der Wagen und Interaktionen  In unterschiedlichen Feldphasen (Marquardt und West 2016, S. 29)

	<p>platz, Stadtsportwagen, mobiler Fairteiler, mobile Kommunikationsblase, mobiles Beet, mobiler Solarwagen, Fragezeichenwagen) (Marquardt und West 2016, S. 29)</p> <p>Allgemein: Evaluation von transdisziplinären Forschungs- und Kommunikationsformaten (Defila und Di Giulio 2018c, S. 388)</p>	<p>Netzwerkstrukturen und Optimierungsmöglichkeit, Transferierbarkeit des Konzepts auf andere Projekte ermöglichen (Geographisches Institut Heidelberg 2018);</p>	<p>zwischen Passant*innen, Neugierigen, Interagierenden und den Wissensorten, d.h. nutzende Akteure und Feld im öffentlichen Raum (West 2018, S. 354)</p> <p>Evaluation der <b>Gespräche und Beobachtungen</b>, wie die „Konfrontierten“ auf die Intervention reagieren (als Hinweis auf mittel- und längerfristige Transformationspotential) (ebd., S. 363)</p> <p>Evaluation auf Basis der <b>multidisziplinär-orientierten Reflexion</b> (als Einzel- oder Gruppenbericht), die theoretische Grundlagen, Erkenntnistheorie, Methodenspektrum und die erfassten Wirkungen auf individueller und gesellschaftlicher Ebene einschließt (ebd., S. 369)</p>
<p><b>EnSign</b></p>			
<p><b>Space Sharing</b></p>			
<p><b>BUGA:log</b></p>	<p>Automatisiertes Transportsystem im öffentlichen Raum, das im Reallabor entwickelt wird</p>		<p><b>Mixed-Method-Ansatz</b> (Marsden et al. 2018, S. 4) Befragung mittels <b>Fragebogen</b> (ebd., S. 5)</p> <p>Abschließende Evaluation in einem iterativen Forschungs- und Entwicklungsprozess</p>

	(Marsden et al. 2018, S. 4)			(Marsden et al. 2019, S. 29)
<b>Reallabor Schorndorf</b>	Quartierbussystem in der Pilotphase (Schnieder und Gebhardt 2016, S. 7)	Weiterentwicklung des Systems bzw. Prototypens ; Verifizierung und Validierung des Systems (König et al. 2016, S. 5)	<p><b>Wiederholte Evaluation von Ergebnissen:</b> von jedem Schritt ausgehende Rückschleifen in vorherige Schritte, um die Ergebnisse miteinander abzugleichen (ebd., S. 5)</p> <p><b>Formative Evaluation durch Partizipation (Bürgerdialog, Test-Nutzer*innen, Fokusgruppen und Befragungen),</b> um Gestaltungsentwürfe und Prototypen des Labordemonstrators zu evaluieren. Davon ausgehend sollen Verbesserungsvorschläge abgeleitet werden und durch kontinuierliche Partizipation wiederholt evaluiert werden (ebd., S. 5)</p> <p><b>GOMS-Methode</b> (Zerlegung einer Aufgabe in Ziele und Teilziele: goals, notwendige Operationen: operators, Methoden: methods und Selektionsregeln bei konkurrierenden Methoden: selction rules) zur Evaluation der Benutzungsoberflächen, Analyse möglicher Bedienstrategien und Vorhersage der Effektivität und Effizienz der Interaktion. Damit erste Möglichkeit zur Evaluation der Usability des Systems (König et al. 2016, S. 6)</p>	Wiederholte Evaluation von Zwischenergebnissen (ebd., S. 5)

				<p><b>Methoden der Usability Evaluation:</b>  <b>Thinking Aloud und Usability-Fragebogen</b>, der auf den sieben Grundsätzen der ISO-Norm 9241-110 zur Ergonomie der Mensch-System-Interaktion basiert (ebd., S. 6)</p>
<b>GO Karlsruhe</b>				<p>Das GO-Karlsruhe Team          (Häußler et al. 2019, S. 397)</p> <p><b>Vorher-Nachher-Vergleich</b> (ebd., S. 397)</p>
<b>Energielabor Tübingen</b>				
<b>Stadt:quartiere 4.0</b>	<p>die verwendeten digitalen und analogen Werkzeuge im Rahmen von „Kinderbeteiligung Gaisburg“(Städtebau-Institut Stuttgart 2019, S. 27)</p>	<p>Evaluation der Werkzeuge in Bezug auf die Eignung für die Beteiligung von Kindern (ebd., S. 27)</p>	<p>Funktionalität, Aufwand und Qualität der Ergebnisse (ebd., S. 27)</p>	
<b>STADT-RAUM-BILDUNG</b>				

<b>Reallabor Asylsu- chende in der Rhein- Neckar-Re- gion</b>				
<b>WTW</b>		Co-Evaluation	<b>Forschungstagebücher</b> (Rose et al. 2019, S. 11) <b>Sitzungsprotokolle</b> (ebd., S. 11) <b>Wirkungsabschätzungsworkshops</b> (ebd., S. 11) <b>Storywallworkshops</b> (ebd., S. 11) <b>Zwischenevaluation mittels strukturierter Interviews</b> (ebd., S. 11) <b>Entwicklung von Leitfäden</b> (ebd., S. 12) <b>Interne Evaluationsworkshops zur Formulierung der Kernergebnisse</b> (ebd., S. 12)	
<b>WM</b>		Co-Evaluation	<b>Befragungen der Bewohner*innen</b> (ebd., S. 11) <b>Memos</b> (ebd., S. 11) <b>Nutzungsbefragung</b> (ebd., S. 11) <b>Prozessbewertung mit Praxispartner*innen und Wissenschaftler*innen</b> (ebd., S. 11) <b>Integrationsworkshops</b> , um (Zwischen-) Ergebnisse zu diskutieren und zu bewerten (Rose et al. 2019, S. 12)	vor, während und nach der Durchführungsphase
<b>KinChem</b>		Co-Evaluation, Ergebnisaufbereitung	<b>Protokolle auf Basis von Audio- und Fotodokumentationen</b> (ebd., S. 11) <b>Befragungen</b> (ebd., S. 11)	



durch Master-  
studierende

**Tab. 6: Sortierung der verwendeten Methoden in den betrachteten Reallaboren nach ihrem Partizipationsgrad.** In Abhängigkeit davon wie viel bzw. in welcher Form Evaluierende und Evaluierte im Rahmen einer bestimmten Evaluationsmethode partizipiert haben, wurde hier nach partizipativen, beobachtenden und befragenden Methoden unterschieden. Nicht zuordenbare Ansätze wurden als weitere Methoden vermerkt (eigene Darstellung).

Partizipative Methoden	Beobachtende Methoden	befragende Methoden	weitere Methoden
<ul style="list-style-type: none"> <li>- Gespräche</li> <li>- Bürgerdialog</li> <li>- Test-Nutzer*innen</li> <li>- Fokusgruppe</li> <li>- Wirkungsabschätzungsworkshop</li> <li>- Storywallworkshop</li> <li>- Prozessbewertung</li> </ul>	<ul style="list-style-type: none"> <li>- Raum- und Forschungstagebuch</li> <li>- Beobachtungen</li> <li>- Inhaltsanalyse</li> <li>- GOMS-Methode</li> <li>- Protokolle</li> <li>- Memos</li> </ul>	<ul style="list-style-type: none"> <li>- Interviews</li> <li>- Befragungen</li> <li>- Feedbackkarten</li> <li>- Fragebogen</li> <li>- Usability Evaluation (Thinking Aloud und Usability-Fragebogen)</li> </ul>	<ul style="list-style-type: none"> <li>- Prä-Post-Design</li> <li>- Multidisziplinär-orientierte Reflexion</li> <li>- Mixed-Method-Ansatz</li> <li>- Vorher-Nachher-Vergleich</li> </ul>

## 5.3 Was lernen wir daraus?

### Welche Evaluationsansätze in Reallaboren gibt es bereits?

Wie sich im Laufe der Literaturrecherche gezeigt hat, ist die Informationslage hinsichtlich möglicher Evaluationsansätze von Reallaboren noch sehr verteilt. Es gibt noch keine einheitlichen Leitlinien, lediglich erste divergierende Praxiserfahrungen oder konzeptionelle Ansatzpunkte. Konkret sind dies insbesondere:

- die Co-Evaluation als Teil eines zirkulären Reallabormodells nach dem Modell des Wuppertal-Instituts
- die begleitende externe Evaluation durch ein Evaluationsgremium, nach dem Vorbild der BaWü-Labs
- konzeptionelle Entwürfe, wie sie in 5.1.3 nach dem jeweiligen Maßstab des Evaluationsgegenstandes gegliedert wurden
- interne Evaluationen durch das Reallaborteam, wie in den analysierten BaWü-Labs

Die jeweiligen Ansätze unterscheiden sich insbesondere im Partizipationsgrad der Reallaborakteure bzw. hinsichtlich der Wahl der Evaluatormethoden, d.h. der Methodik der Datensammlung und -auswertung ist ebenfalls vielseitig, wobei auch hier zwischen verschiedenen Graden der Partizipation unterschieden werden kann, wie in Tab. 6 gezeigt wurde. Mit dem Grad der Partizipation in der Evaluation geht letztlich auch ein unterschiedlicher (personeller, zeitlicher und damit kostenbezogener) Aufwand einher, welcher in der Konzeption beachtet werden muss. Die hier gesammelten Ansätze bilden eine erste Sammlung und Sortierung bisheriger Konzepte, an welche nach den Bedarfen der Evaluation angeknüpft werden kann.

### Was sollte in der zukünftigen Forschung beachtet werden?

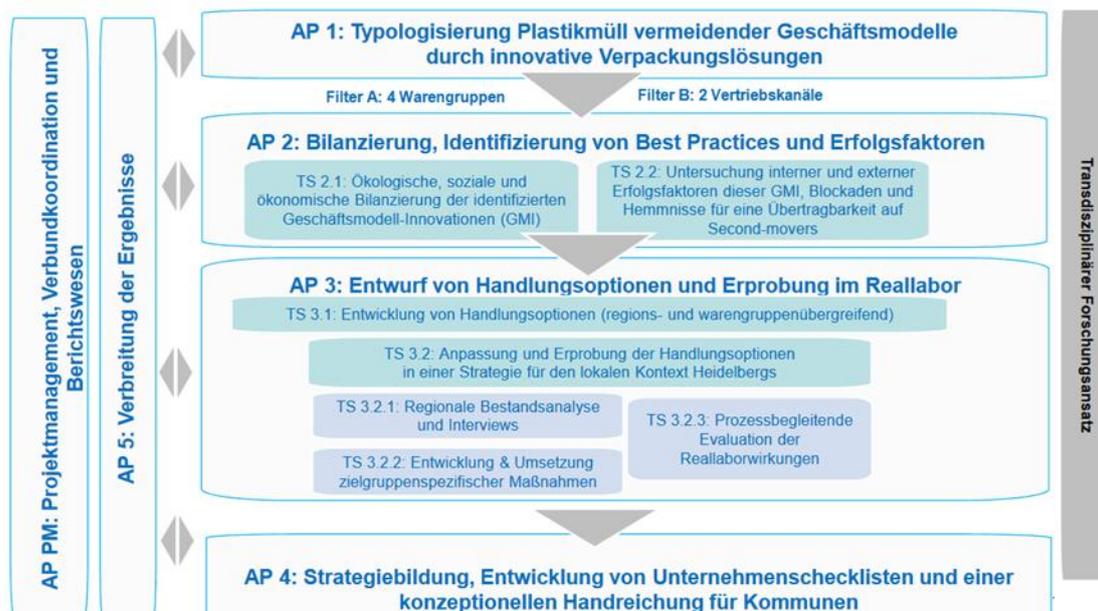
Für die weitere Evaluationsforschung und -praxis in Reallaboren ist es notwendig mehr Transparenz über bisherige Evaluationen zu schaffen. Dies könnte ermöglichen den transformativen Charakter der Reallabore besser bewerten zu können bzw. die Methoden für die Evaluierung passgenauer auswählen zu können für nachfolgende Vorhaben. Dieser Lerneffekt könnte dadurch verstetigt werden, wenn eine zentrale Sammlung der Ansätze erfolgt. Die hier vorgestellte, aufgeschlüsselte Gliederung nach Gegenstand, Zweck, Kriterien, Evaluatormethoden, Methoden und Zeitpunkt kann dazu ein hilfreiches Muster sein. In einem weiteren Schritt könnten auch Metaevaluationen, das heißt die Evaluation der Evaluationsansätze angedacht werden, um die Praxistauglichkeit der Evaluationen zu bewerten.

## 6 Übertrag Innoredux

### 6.1 Einbettung von Reallabor und Evaluation in das Gesamtprojekt

Das Reallabor ist der zentrale Baustein des Forschungsprojekts Innoredux. Im formalen Projektablauf ist es im dritten Arbeitspaket (AP 3) verordnet, wie auf Abb. 8 zu sehen ist. Damit schließt es an die zwei vorherigen Arbeitspakete an, in welchen zu Beginn des Reallabors bereits Geschäftsmodelle typologisiert, Verpackungslösungen bilanziert sowie Best Practices und Erfolgsfaktoren identifiziert sein wurden. Im Reallabor sollen, an diese Arbeitspakete anschließend, Lösungsansätze bzw. Geschäftsmodelle entwickelt, untersucht und in der Praxis erprobt werden (BMBF 2019). Zunächst soll gemeinsam mit den beteiligten Praxisakteuren in einem Co-Design-Prozess Zielsetzung und Absicht der Realexperimente und Fragestellungen definiert werden. Darauf aufbauend werden Handlungsoptionen, im Sinne von Verpackungsoptionen und Unterstützungsmöglichkeiten, entwickelt (siehe Abb. 10). Diese Handlungsoptionen sind zunächst regions- und warengruppenübergreifend, werden anschließend jedoch in eine Strategie für den lokalen Kontext in Heidelberg angepasst (TS 3.2). Dies geschieht mit Hilfe einer regionalen Bestandsanalyse und ergänzenden Interviews (TS 3.2.1). So können daraus zielgruppenspezifische Maßnahmen entwickelt und umgesetzt werden (TS 3.2.2). Die Ergebnisse des Reallabors sollen anschließend, gemeinsam mit den Ergebnissen der vorherigen Arbeitspakete in die Entwicklung von Unternehmenschecklisten und einer konzeptionellen Handreichung für Kommunen münden (AP 4). Die Evaluation erfolgt prozessbegleitend im dritten Arbeitspaket, im Rahmen des Reallabors, unter TS 3.2.3.

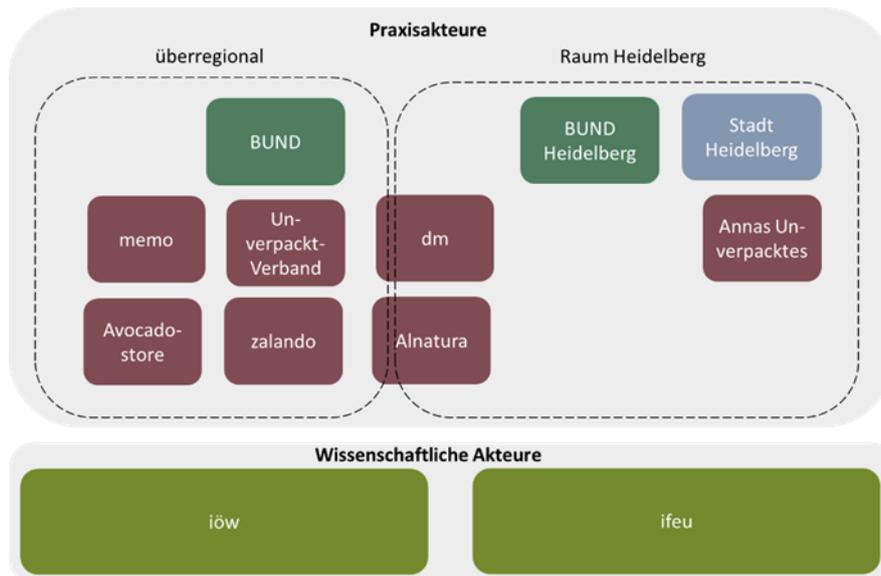
Der Begriff „intraprojektdynamisch entstehendes Reallabor“ wurde deshalb gewählt, da sich das Reallabor erst aus dem Kontext des Projektes heraus entwickelt. Damit unterscheidet sich Innoredux beispielsweise von den BaWü-Labs, da diese von Anfang an als Reallabore konzipiert wurden, wohingegen das Reallabor im Projekt Innoredux in eine, das Reallabor überlagernde, Projektstruktur eingebettet ist. Inwiefern diese Reallabor-konstruktion Auswirkungen hat auf Rollenverteilungen und -veränderungen, projektinterne Emanzipations- und Kommunikationsprozesse sowie weitere projekt- und reallaborformende Strukturen kann erst im Durchführungsprozess beobachtet werden.



**Abb. 8: Projekttablauf von Innoredux.** Das Projekt gliedert sich in fünf aufeinander aufbauende Arbeitspakete (AP 1-5) sowie ein übergreifendes, welches Projektmanagementtätigkeit zur Aufgabe hat (AP PM). Die jeweiligen Aufgaben bzw. Zielsetzungen sind in der Abbildung in den jeweiligen Kästchen dargestellt. Das Reallabor ist zentral im dritten Arbeitspaket angelegt (AP 3).

## 6.2 Akteurskonstellation

Nachdem es sich bei Innoredux um ein transdisziplinäres Projekt handelt, arbeiten auch dort Akteure aus Wissenschaft und Praxis zusammen. Die wissenschaftliche Akteursseite wird aus Mitarbeitenden des IÖW und des ifeu gebildet (Webseite Innoredux o.J.). Bei den Praxispartnern handelt es sich um Akteure aus dem Einzelhandel (memo, Zalando, Avocado-store, Unverpacktverband, dm, Alnatura, Annas Unverpacktes), kommunalen Akteure der Stadt Heidelberg sowie Akteure aus der überregionalen und regionalen Ortsgruppe des Umwelt- und Naturschutzbundes BUND (ebd.). Alle beteiligten Akteure werden in Abb. 9 auf Organisationszugehörigkeitsebene visualisiert. Nachdem es sich bei Reallaboren um ein Forschungsformat mit geographischem Bezug handelt, wurde es als relevant erachtet zwischen regionalen und überregionalen Praxisakteuren zu differenzieren. Diejenigen Einzelhandelsakteure, die nicht ausschließlich im Raum Heidelberg operieren, dort aber ebenfalls Filialen haben, wurden zwischen diese Kategorisierungspole gesetzt. Die wissenschaftlichen Akteure wurden nicht nach ihrer geographischen Zugehörigkeit differenziert, da sowohl IÖW, als auch ifeu-seitig Akteure in Heidelberg ansässig und mit den örtlichen Gegebenheiten vertraut sind. Alle dargestellten Akteure sind im Rahmen der Reallabordurchführung letztlich auch von der prozessbegleitenden Evaluation betroffen und sollten deshalb Beachtung finden.



**Abb. 9: Akteurskonstellation im Projekt Innoredux.** Die Abbildung zeigt die transdisziplinäre Zusammensetzung der wissenschaftlichen Akteure (hellgrün), Akteure aus dem Einzelhandel (rot), Akteuren aus den Umweltverbänden (dunkelgrün) und dem kommunalen Akteur (grau). Die Praxisakteure wurden nach ihrem geographischen Bezugspunkt geordnet, da das Reallabor des Projekts selbst in Heidelberg situiert sein wird, jedoch nicht alle Akteure dort angesiedelt sind. Diejenigen Einzelhandelsakteure die überregional agieren, aber Filialen in Heidelberg haben, wurden zwischen diese Pole gesetzt (eigene Darstellung).

## 6.3 Konzeptionelle Anknüpfungspunkte

In früheren Kapiteln wurden bereits verschiedene bisherige Evaluationsansätze beschrieben und sortiert. Um diese für die Erstellung eines Evaluationsschemas für das Reallabor in Innoredux anwendbar zu gestalten, wurden erste Empfehlungen/Leitgedanken formuliert (Tab. 7) sowie eine Überblicksmatrix erstellt (Tab. 8). Diese sollen als konzeptionelle Anknüpfungspunkte an die bisherige Forschung dienen und den Weg für die konkrete Ausformulierung des Evaluationsschemas bereiten.

Tab. 7 beruht auf den von Stockmann und Meyer (2014, S. 75) formulierten und bereits in der Auswertung der Evaluationspraxis der BaWü-Labs angewendeten, Leitfragen „Was (welcher Gegenstand) wird wozu (zu welchem (Zweck) anhand welcher Kriterien von wem wie (mit welchen Methoden) evaluiert?“. Diese wurden in modifizierter Form den Empfehlungen/Leitgedanken zugeordnet, mit dem Ziel Impulse und Anregungen für die Evaluationskonzeption zu bieten. Mit dieser Form der Darstellung soll bezweckt werden, dass relevante Diskussionspunkte für die Konzeptionierung schnell und einfach auffindbar sind.

In Tab. 8 findet sich eine abschließende Zusammenschau mit möglichen Auswahloptionen zur Entwicklung des Evaluationsdesigns. Diese wurde konzipiert, um eine Übersicht der Evaluationsmöglichkeiten in Reallaboren zu schaffen und darauf aufbauend die konkrete Ausgestaltung der Evaluation des Innoredux-Reallabors zu ermöglichen, aber auch im Gesamtkontext zu verorten. Auch in dieser Matrix finden sich die Leitfragen in den Spalten wie-

der. Dabei sind den verschiedenen, aus der Literatur entnommenen, Optionen jeweils Indizes zugeteilt. So sind den Evaluationsgegenständen die Indizes G1-G3, den Zweckoptionen Z1-Z5, den Auswahlmöglichkeiten für die Kriterien K1-K3, der Methodenauswahl M1-M3, der Auswahl der Evaluierenden E1-E4 und dem Zeitpunkt P1-4 zugeordnet. Bei der Evaluationskonzeption kann so nach einem Codesystem vorgegangen werden. Das bedeutet, dass je nachdem welche Evaluationsoption für die jeweilige Spalte gewählt wird, sich für die Gesamtevaluation ein bestimmter Code ergibt. So kann beispielsweise für die Evaluation eines Realexperiments (G2), mit dem Zweck der Optimierungsfunktion (Z4), mit der Wahl des Kriterienkatalogs von Luederitz et al. (2017) (K2), der Wahl von partizipativen Methoden, im Kontext einer Co-Evaluation (E4), durch iterative Evaluationsvorgänge (P4) der Code G2-Z4-K2-E4-P4 zugewiesen werden. Damit lässt sich einerseits die Evaluationskonzeption in Kurzform beschreiben, aber auch in der bisherigen Evaluationslandschaft von Reallaboren verorten. Die Matrix kann in ihren Zeilen erweitert werden, wenn hier nicht beachtete Optionen eingetragen werden.

**Tab. 7: Leitfragen und komprimierte Übersicht der Empfehlungen und Leitgedanken.** Mit dieser Tabelle sollen die zuvor diskutierten Punkte in eine übersichtliche Darstellung gebracht werden, um so konkret und anschaulich Impulse für die Evaluationskonzipierung bieten zu können (eigene Darstellung).

Leitfragen	Empfehlungen/ Leitgedanken für die Konzeptionierung
Auf welcher Maßstabsebene soll evaluiert werden/ Welcher Evaluationsgegenstand ist empfehlenswert?	Evaluation auf Ebene der Realexperimente/Maßnahmen
	<p>Es kann auf ein fertiges, von Luederitz et al. (2017) entwickeltes Evaluationsschema zurückgegriffen werden.</p> <p>Im Projekt bieten sich Anknüpfungsmöglichkeiten an bereits existierendes Wissen über verpackungsmindernde Maßnahmen. Dadurch sind eine Übertragbarkeit und Kontextualisierung der Evaluationsergebnisse ggf. einfacher zu realisieren.</p>
	Reallaborübergreifende Evaluation
	<p>Es ist kein ausgereifter Kriterienkatalog vorhanden, woraus sich in der Konzeption ein höherer Aufwand ergibt.</p> <p>Mit dem Reflexionsschema von Beecroft et al. (2018) existiert jedoch eine Grundlage, die für die Entwicklung partizipativer Evaluationsprozesse genutzt werden kann.</p>
	Evaluation der Methoden
	<p>In Kombination mit einer reallaborübergreifenden Evaluation ist eine Evaluation der Methoden denkbar.</p> <p>Auch hier entsteht bei einer potentiellen Umsetzung ein höherer konzeptioneller Aufwand, allerdings könnte auf das Reflexionsschema von Beecroft et al. (2018) und die Gütekriterien von Defila und Di Giulio (2018) zurückgegriffen werden, um ein passendes Evaluationsdesign zu entwickeln.</p>

Zu welchem Zweck wird evaluiert?	<p>Es wird empfohlen im Co-Design-Prozess die Rolle der Evaluation im Reallabor zu eruieren. Denn vorrangige Reflexionen zum gewünschten Zweck der Evaluation sind essentiell, da Klarheit über die Funktion der Evaluation eine größtmögliche Nutzbarkeit der Ergebnisse ermöglicht und Erwartungen diesbezüglich abgeklärt werden können.</p> <p>Die in Abb. 12 angeführten Leitfragen können eine Grundlage für diese Überlegungen darstellen.</p>
Welche Kriterien können/sollen für die Evaluation genutzt werden?	<p>Gütekriterien von Defila und Di Giulio (2018a, siehe Anhang, Tab. 11)</p> <p>Ursprünglich waren diese zwar nicht für Evaluationszwecke konzipiert, sie sind jedoch ggf. für die Qualitätsevaluation der einzusetzenden oder eingesetzten Methoden operationalisierbar.</p> <hr/> <p>Kriterienkatalog von Luederitz et al. (2017, siehe Anhang, Tab. 10)</p> <p>Der Kriterienkatalog ist für die Evaluation von Realexperimenten ausgelegt, bereits einsatzfähig und daher bei der entsprechenden Maßstabsebene zu empfehlen.</p>
	<p>Partizipative Entwicklung von Kriterien</p> <p>Durch diesen Ansatz können Perspektiven und Zielvorstellungen der verschiedenen Akteure bestmöglich berücksichtigt werden, woraus sich eine hohe Nutzbarkeit für die Beteiligten ergeben kann.</p>
Welche Methoden eignen sich für die Evaluation?	<p>Hierzu können keine klaren Empfehlungen gegeben werden, da dies stark vom Zweck der Evaluation und dem betrachteten Evaluationsgegenstand abhängt.</p> <p>Sobald andere Leitfragen geklärt wurden, können die in Tab. 6 nach Partizipationsgrad gegliederten Methoden für die Auswahl genutzt werden.</p>
Wer soll die Aufgabe der Evaluation übernehmen?	<p>Wenn es sich um eine Legitimationsfunktion der Evaluation handelt, wird empfohlen eine interne Fremdevaluation durchzuführen.</p> <p>Ansonsten sollten zunächst die zeitlichen und personellen Ressourcen sowie die Motivation aller Akteure berücksichtigt werden, um in Abhängigkeit davon den Grad der Partizipation zu bestimmen. Möglichkeiten der Partizipation bieten die Co-Interpretation der Ergebnisse oder alternativ eine gemeinsame Entwicklung der Evaluationskriterien.</p> <p>Ansonsten kann die Evaluation als Selbst-Evaluation durch Beteiligte des Reallabors erfolgen.</p>

---

Wann und wie oft  
soll evaluiert wer-  
den?

Es wird angeraten mehrfache Evaluationsdurchgänge durchzuführen, so wie es auch von Luederitz et al. (2017) empfohlen wird sowie teilweise in den untersuchten BaWü-Labs durchgeführt wurde.

---

**Tab. 8: Abschließende Zusammenschau mit möglichen Auswahloptionen zur Entwicklung des Evaluationsdesigns.** Das hier dargelegte Schema soll eine Hilfestellung für die Evaluationskonzipierung des Reallabors in Innoredux bieten, indem verschiedene, in der Arbeit angesprochene, Ansätze zusammenfassend gegliedert werden (eigene Darstellung).

<b>Gegenstand der Evaluation</b> (vgl. Kapitel 4.1, Kapitel 5.1.3)	<b>Zweck der Evaluation</b> (vgl. Kapitel 4.2)	<b>Kriterienwahl</b> (vgl. Kapitel 4.3)	<b>Methodenwahl</b> (vgl. Kapitel 5, Tab. 6)	<b>Evaluierende</b> (vgl. Kapitel Fehler! Verweisquelle konnte nicht gefunden werden.4)	<b>Evaluationszeitpunkt</b>
<b>G1</b> Reallaborübergreifende Evaluation	<b>Z1</b> Legitimationsfunktion	<b>K1</b> Gütekriterien von Defila und Di Giulio (2018a)  (s. Anhang, Tab. 11)	<b>M1</b> Partizipative Methoden	<b>E1</b> externe Evaluation (Evaluationsgremium)  (s. Kapitel 5.1.2)	<b>P1</b> Ex-ante
<b>G2</b> Evaluation der Realexperimente	<b>Z2</b> Lern- und Dialogfunktion	<b>K2</b> Kriterienkatalog von Luederitz et al. (2017)  (s. Anhang, Tab. 10)	<b>M2</b> Beobachtende Methoden	<b>E2</b> Interne Fremdevaluation	<b>P2</b> formativ
<b>G3</b> Evaluation der Methoden	<b>Z3</b> Erkenntnisfunktion	<b>K3</b> Partizipative Entwicklung der Kriterien	<b>M3</b> Nachfragende Methoden	<b>E3</b> Evaluation durch das Forschungsteam (Selbst-Evaluation)  (vgl. BaWü-Labs in Tab. 3)	<b>P3</b> Ex-post
	<b>Z4</b> Optimierungsfunktion			<b>E4</b> Co-Evaluation  (s. Kapitel 5.1.1)	<b>P4</b> Iterativ  (vgl. BaWü-Labs in Tab. 3)
	<b>Z5</b> Entscheidungsfunktion				

## 7 Fazit

Mithilfe der vorliegenden Untersuchung sollte eine Antwort auf die Frage gefunden werden, wie Reallabore bisher evaluiert wurden. Dazu wurden in dieser Arbeit die bisherigen Ansätze zur Evaluation von Reallaboren gesammelt und systematisiert. Dabei stachen zwei Hauptansätze heraus: Die externe Evaluation durch ein Evaluationsgremium nach dem Beispiel der BaWü-Labs sowie die Co-Evaluation in einem zirkulären Projektansatz. Für die weiteren Ansätze wurde eine Systematisierung nach der Maßstabsebene des betrachteten Evaluationsgegenstandes vorgenommen, wobei nach reallaborübergreifender Evaluation, Evaluation von Realexperimenten sowie der Evaluation von Methoden gegliedert wurde.

Mittels der durchgeführten qualitativen Literaturrecherche konnte zudem die Evaluationspraxis in den BaWü-Labs der beiden Förderlinien näher untersucht werden. Dazu wurden die gefundenen Informationen nach Evaluationsgegenstand, Zweck, Kriterienauswahl, Methodenauswahl, Evaluierenden und Evaluationszeitpunkt strukturiert. Zentrale Ergebnisse waren, dass vorwiegend umgesetzte Maßnahmen bzw. Realexperimente evaluiert wurden. Hierzu wurden teilweise Erfolgskriterien genutzt, wobei insgesamt wenig Informationen zur Kriterienauswahl vorliegen. Zudem wurden Methoden unterschiedlicher Partizipationsgrade zwischen Evaluierten und Evaluierenden genutzt, wobei nach partizipierenden, nachfragenden und beobachtenden Methoden unterschieden werden kann. Auch wurde gezeigt, dass Evaluation in den BaWü-Labs nicht nur extern durch das Evaluationsgremium stattgefunden hat, sondern ebenso intern in den Reallaboren selbst. Die Evaluation erfolgte dabei oftmals in iterativen oder mehrmaligen Evaluationsdurchläufen.

Damit bietet die vorliegende Arbeit zum ersten Mal einen synoptischen Blick zum bisherigen Vorgehen zur Evaluierung von Reallaboren, insbesondere bzgl. der BaWü-Labs.

Ausgehend von dem Überblick über die diversen Evaluationskonzeptionen und den Erkenntnissen aus den BaWü-Labs wurden Praxisempfehlungen für ein intraprojekt-dynamisch entstehendes Reallabor im Projekt Innoredux abgeleitet. Damit sollte der Frage nachgegangen werden, welche Methodik für die Anwendung in einem intraprojekt-dynamisch entstehenden Reallabor geeignet sein kann. Nachdem das Reallabor erst noch aufgebaut werden muss, konnten hierbei jedoch lediglich erste praxisorientierte Leitgedanken und Auswahlmöglichkeiten formuliert werden. Diese wurden einerseits ausführlich diskutiert jedoch auch in zwei komprimierten Darstellungen für eine Praxisnutzung operationalisiert. Als konkretes Ergebnis wurde eine tabellarische Übersicht mit Empfehlungen erstellt und eine Zusammenschau möglicher Vorgehensweisen zur Evaluation des Innoredux-Reallabors generiert.

## 8 Literaturverzeichnis

- Arnold, Annika; Piontek, Felix M. (2018): Zentrale Begriffe im Kontext der Reallaborforschung. In: *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden, S. 143–154.
- Beecroft, Richard; Trenks, Helena; Rhodius, Regina; Benighaus, Christina; Parodi, Oliver (2018): Reallabore als Rahmen transformativer und transdisziplinärer Forschung: Ziele und Designprinzipien. In: *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden, S. 75–100.
- Bergmann, Matthias; Brohmann, Bettina; Hoffmann, Esther; Loibl, Marie Céline; Rehaag, Regine; Schramm, Engelbert et al. (2005): Qualitätskriterien transdisziplinärer Forschung: ein Leitfaden für die formative Evaluation von Forschungsprojekten: ISOE.
- BMBF (2019): Innoredux. Geschäftsmodelle zur Reduktion von Plastikmüll entlang der Wertschöpfungskette: Wege zu innovativen Trends im Handel. Online verfügbar unter <https://bmbf-plastik.de/de/verbundprojekt/innoredux>, zuletzt geprüft am 14.03.2020.
- Borner, Joachim; Kraft, Anne H. (2018): Konzeptpapier zur Reallabor-Methode im ENavi-Projekt. In: *Berlin: KMGNE*. <https://doi.org/10.13140/RG.2.10639.28322>, S. 1.
- Defila, Rico; Di Giulio, Antonietta (2018a): Partizipative Wissenserzeugung und Wissenschaftlichkeit—ein methodologischer Beitrag. In: *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden, S. 39–67.
- Defila, Rico; Di Giulio, Antonietta (2018b): Reallabore als Quelle für die Methodik transdisziplinären und transformativen Forschens—eine Einführung. In: Rico Defila und Antonietta Di Giulio (Hg.): *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden, S. 9–35.
- Defila, Rico; Di Giulio, Antonietta (Hg.) (2018c): *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden.
- Defila, Rico; Di Giulio, Antonietta (2018d): What is it good for? Reflecting and systematizing accompanying research to research programs. In: *GAIA-Ecological Perspectives for Science and Society* 27 (1), S. 97–104.
- Defila, Rico; Di Giulio, Antonietta (Hg.) (2019a): *Transdisziplinär und transformativ forschen*, Band 2. Eine Methodensammlung. Springer Fachmedien Wiesbaden GmbH
- Defila, Rico; Di Giulio, Antonietta (2019b): Wie Reallabore für Herausforderungen und Expertise in der Gestaltung transdisziplinären und transformativen Forschens sensibilisieren—eine Einführung. In: *Transdisziplinär und transformativ forschen*, Band 2: Springer, S. 1–30.
- DeGEval (2016): *Standards für Evaluation*. Erste Revision 2016. 1. Aufl. Mainz: DeGEval - Gesellschaft für Evaluation
- Döring, Nicola; Bortz, Jürgen (2016): *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5. Aufl.: Berlin Heidelberg: Springer.
- Geiger, Sonja; Hirscher, Anja; Müller, Martin (2017): Maßnahmenevaluation im transdisziplinären Forschungssetting. In: *GAIA-Ecological Perspectives for Science and Society* 26 (2), S. 147–148.
- Geographisches Institut Heidelberg (2018): Urban Office. Online verfügbar unter [https://www.geog.uni-heidelberg.de/hca/reallabor\\_tp2.html](https://www.geog.uni-heidelberg.de/hca/reallabor_tp2.html), zuletzt geprüft am 18.05.2020.
- Häußler, Elke; Blaszczyk, Robert; Eckart, Jochen; Hupfer, Christoph (2019): Den Fußverkehr in Städten fördern: Partizipative Forschung im Reallabor GO Karlsruhe. In: *GAIA-Ecological Perspectives for Science and Society* 28 (4), S. 396–397.

ifeu (2019): Innoredux. Geschäftsmodelle zur Reduktion von Plastikmüll entlang der Wertschöpfungskette: Wege zu innovativen Trends im Handel. Online verfügbar unter <https://www.ifeu.de/projekt/innoredux/>, zuletzt geprüft am 14.03.2020.

KIT (11.04.2019): Reallabore erhöhen Sichtbarkeit. Karlsruher Institut für Technologie, Wuppertal Institut, Leuphana Universität Lüneburg und Ecornet gründen „Netzwerk Reallabore der Nachhaltigkeit“. Online verfügbar unter [http://www.kit.edu/kit/pi\\_2019\\_049\\_reallabore-erhoehen-sichtbarkeit.php](http://www.kit.edu/kit/pi_2019_049_reallabore-erhoehen-sichtbarkeit.php), zuletzt geprüft am 05.04.2020.

König, Alexandra; Brandies, Alexander; Schnieder, Lars; Dotzauer, Mandy; Jipp Meike (Hg.) (2016): Zukunftsszenarien autonomer Fahrzeuge—Nutzungsorientierte Gestaltung eines individuell abrufbaren Personentransportsystems. Proceedings of the 5 Interdisziplinärer Workshop Kognitive Systeme, Bochum, Germany.

Lazarova, Kristin; Helfenstein, Basil S.; Dietz, R.; Alcántara, S. (2018): Parklets für Stuttgart. Ein Realexperiment. In: *Stuttgart: Reallabor für nachhaltige Mobilitätskultur Universität Stuttgart*. [http://www.rnm.net/wp-content/uploads/2018/03/Parklets-fuer-Stuttgart\\_web.pdf](http://www.rnm.net/wp-content/uploads/2018/03/Parklets-fuer-Stuttgart_web.pdf). Zugegriffen am 27.

Luederitz, Christopher; Schöpke, Niko; Wiek, Arnim; Lang, Daniel J.; Bergmann, Matthias; Bos, Joannette J. et al. (2017): Learning through evaluation—A tentative evaluative scheme for sustainability transition experiments. In: *Journal of Cleaner Production* 169, S. 61–76.

Marquardt, Editha; West, Christina (2016): Co-Produktion von Wissen in der Stadt. In: *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 25 (3), S. 26–31.

Marsden, Nicola; Bernecker, Tobias; Zöllner, Raoul; Sußmann, Nico; Kapser, Sebastian (Hg.) (2018): BUGA: log—A Real-World Laboratory Approach to Designing an Automated Transport System for Goods in Urban Areas. 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC): IEEE.

Marsden, Nicola; Dierolf, Nicole; Herling, Claudia (2019): Partizipative und sozialverantwortliche Technikentwicklung auf der letzten Meile. In: *Mensch und Computer 2019-Workshopband*.

NaWiKo (o.J.): Konsumverhalten und Innovationen zur nachhaltigen Chemie am Beispiel von Produkten mit problematischen Inhaltsstoffen (KInChem). Online verfügbar unter <https://nachhaltigeswirtschaften-soef.de/kinchem>, zuletzt geprüft am 17.06.2020.

Parodi, Oliver; Ley, Astrid; Fokdal, Josefine; Seebacher, Andreas. (2018): Empfehlungen für die Förderung und den Aufbau von Reallaboren. Positionspapier der BaWü-Labs.

Rose, Michael; Wanner, Matthias; Hilger, Annaliesa (2019): Das Reallabor als Forschungsprozess und-infrastruktur für nachhaltige Entwicklung: Konzepte, Herausforderungen und Empfehlungen. Wuppertal: Wuppertal Institut für Klima, Umwelt, Energie.

Schöpke, Niko; Stelzer, Franziska; Caniglia, Guido; Bergmann, Matthias; Wanner, Matthias; Singer-Brodowski, Mandy et al. (2018): Jointly experimenting for transformation? Shaping real-world laboratories by comparing them. In: *GAIA-Ecological Perspectives for Science and Society* 27 (1), S. 85–96.

Schnieder, Lars; Gebhardt, Laura (2016): Nutzerorientierter Entwurf innovativer Mobilitätskonzepte für urbane Räume. In: *14. Fachtagung EKA 2016—Entwurf komplexer Automatisierungssysteme*.

Schönau, Manuela; Kasten, Peter; Birzle-Harder, Barbara; Kurzrock, Björn-Martin; Rubik, Frieder; Deffner, Jutta (2018): Nachhaltigkeitswirkungen wohnbegleitender Dienstleistungen in gemeinschaftlichen Wohnformen. Analyse von drei Praxisbeispielen gemeinschaftlicher Flächennutzung und Mobilitätsangebote. Berlin. Online verfügbar unter [http://www.wohnmobil-projekt.de/fileadmin/user\\_upload/Downloads/Nachhaltigkeitswirkungen\\_Werkstattbericht.pdf](http://www.wohnmobil-projekt.de/fileadmin/user_upload/Downloads/Nachhaltigkeitswirkungen_Werkstattbericht.pdf).

Singer-Brodowski, Mandy; Beecroft, Richard; Parodi, Oliver (2018): Learning in real-world laboratories: A systematic impulse for discussion. In: *GAIA-Ecological Perspectives for Science and Society* 27 (1), S. 23–27.

- Spaapen, Jack; van Drooge, Leonie (2011): Introducing 'productive interactions' in social impact assessment. In: *Research evaluation* 20 (3), S. 211–218.
- Städtebau-Institut Stuttgart (2019): Wandel gestalten. Das Reallabor Stadt:quartiere 4.0. Online verfügbar unter [https://digitale-mitwirkung.de/wp-content/uploads/2019/09/Wandel-gestalten-Das-Reallabor-Stadt-quartiere-4\\_0-2019.pdf](https://digitale-mitwirkung.de/wp-content/uploads/2019/09/Wandel-gestalten-Das-Reallabor-Stadt-quartiere-4_0-2019.pdf), zuletzt geprüft am 18.05.2020.
- Stockmann, Reinhard; Meyer, Wolfgang (2014): Evaluation. Eine Einführung. Stuttgart: UTB Uni-Taschenbücher.
- Unger, Hella von (2014): Partizipative Forschung. Einführung in die Forschungspraxis. Wiesbaden: Springer-Verlag.
- Wagner, Felix (2018): Reallabor Symposium. Online verfügbar unter [https://mwk.baden-wuerttemberg.de/fileadmin/redaktion/m-mwk/intern/dateien/pdf/Forschung/Reallabore/Pr%C3%A4s\\_1\\_Uebersicht.pdf](https://mwk.baden-wuerttemberg.de/fileadmin/redaktion/m-mwk/intern/dateien/pdf/Forschung/Reallabore/Pr%C3%A4s_1_Uebersicht.pdf), zuletzt geprüft am 23.06.2020.
- Wagner, Felix; Grunwald, Armin (2019): Reallabore zwischen Beliebtheit und Beliebigkeit: Eine Bestandsaufnahme des transformativen Formats. In: *GAIA-Ecological Perspectives for Science and Society* 28 (3), S. 260–264. München: Oekom Verlag.
- Wanner, Matthias; Hilger, Annaliesa; Westerkowski, Janina; Rose, Michael; Stelzer, Franziska; Schöpke, Niko (2018): Towards a cyclical concept of real-world laboratories: a transdisciplinary research practice for sustainability transitions. In: *disP-The Planning Review* 54 (2), S. 94–114. Zürich: disP The Planning Review
- Wanner, Matthias; Stelzer, Franziska (2019): Reallabore-Perspektiven für ein Forschungsformat im Aufwind. Wuppertal: Wuppertal Institut für Klima, Umwelt, Energie.
- Webseite Dietenheim zieht an (o. J.): Das Reallabor. Online verfügbar unter <http://dietenheim-zieht-an.de/das-reallabor/>, zuletzt geprüft am 17.06.2020.
- Webseite Innoredux (o. J.): Innoredux - Geschäftsmodelle zur Reduktion von Plastikmüll entlang der Wertschöpfungskette: Wege zu innovativen Trends im Handel - Plastik reduzieren. Online verfügbar unter <https://www.plastik-reduzieren.de/>, zuletzt geprüft am 14.03.2020.
- Webseite WTW (o. J.): Wohlstands-Transformation-Wuppertal. ein urbanes Transformationslabor für nachhaltiges Wirtschaften. Online verfügbar unter <https://w-indikatoren.de/die-transformation/>, zuletzt geprüft am 17.06.2020.
- West, Christina (2018): Wissen to Go –Transdisziplinär-transformative Lehre als „Reallabor im Kleinen. In: *Transdisziplinär und transformativ forschen*: Springer VS, Wiesbaden, S. 329–373.
- Zimpelmann, Beate (2018): Reallabore: Nuklei einer Wissenschaft für Nachhaltigkeit? Reflexionen der Gutachtergruppe zur ersten BaWü-Lab-Förderlinie. In: *GAIA-Ecological Perspectives for Science and Society* 27 (3), S. 321–322.

## 9 Anhang

**Tab. 9: Reflexionsschema.** In diesem Schema werden fünf Designprinzipien von Reallaboren auf die zu erreichenden Zielsetzungen im Durchführungsprozess bezogen. So kann es als Reflexions- und Evaluationsgrundlage in Reallaboren dienen (verändert nach Beecroft et al. 2018, S. 94f.)

Designprinzipien	Forschungsziele	Praxisziele	Bildungsziele
Problem- und Themenangemessenheit herstellen			
Räumliche Angemessenheit herstellen			
Zeitliche Angemessenheit herstellen			
Angemessene Akteursrollen etablieren			
Experimentell-reflexive Arbeitsweise fördern			

**Tab. 10: Schema für die Evaluation von Realexperimenten.** Dargestellt ist ein Evaluationschema, welches es erlaubt Realexperimente nach den Dimensionen Output, Outcome, Process und Input zu evaluieren, wobei diese nicht chronologisch verstanden, sondern als interdependent angesehen werden. Den nach den genannten Dimensionen unterteilten Kriterien sind Erklärungen und Evaluationsfragen zugeordnet. Dem ursprünglichen Schema Luederitz et al. wurden in dieser Tabelle typische Indikatoren hinzugefügt, welche die Autor\*innen im Fließtext genannt haben (Luederitz et al. 2017, S. 64ff.).

Kriterium	Erklärung	Evaluationsfrage	Typische Indikatoren
<b>Output</b>			
<b>Built capacities</b>	Empower participants to act sustainably in everyday decision-making and practices through educating them in cognitive, practical and interper-	Does the transition experiment build capacities in participants to generate sustainability solutions?	Post-experiment activities and practices carried out by participants that have the potential to address the given sustainability problem such as

	sonal competencies and enable to internalize required skills and activate new behavioral patterns.		community gardening and food distribution systems, consumption of organic food products, launching of new sustainability-based businesses, expansion of networks, and incorporation of sustainability into decision making in the public or private sector.
<b>Actionable knowledge</b>	Generate evidence-supported instructions that have been tested on effectively solving a sustainability problem within the defined experimental setting including guidelines on how to most effectively transition from the current to the desired state.	Does the transition experiment generate actionable knowledge that provides evidence on how to generate sustainability solutions?	Scientific output as well as context specific transition pathways that identify strategic actions for implementing transformational change and building agreement on the problem framing.
<b>Accountability</b>	Ensure confidence and commitment of participants to implement results generated by the experiment and their dedication to positive change.	Does the transition experiment build confidence and commitment for generating and realizing sustainability solutions?	Participants' attitudes, but also more formalized commitments towards the implementation of the results.
<b>Structural changes</b>	Changes in physical structures:  Create new or transform existing buildings, infrastructures, technologies and products that are radically different from existing ones.	Changes in physical structures:  Does the transition experiment generate physical changes that support solutions for the identified sustainability problem?	Modified or newly built forms such as new bicycle lanes, rooftops, novel or improved products arising from new scientific knowledge and innovations. Other indicators would be commercialization of patents; shifts in the design, production and manufacturing of

			goods; and changes in the natural environment, for example, afforested areas or increasing green spaces in urban areas
	<p>Changes in societal realms:</p> <p>Create new or transform existing networks and organizations, values and norms, rules and policies, behavior and practices, and discourses that are radically different from existing ones.</p>	<p>Changes in societal realms:</p> <p>Does the transition experiment generate societal changes that support solutions for the identified sustainability problem?</p>	<p>New or altered activities, practices, routines, as well as social relations and partnerships</p>
<b>Facilitate up-take</b>	<p>Transferability:</p> <p>Create generalizable lessons learned regarding processes through to outcome of the experimentation that are applicable to different contexts.</p>	<p>Transferability:</p> <p>Does the transition experiment indicate how the sustainability solution can be transferred to different contexts?</p>	<p>Reliability of insights in other contexts or validity of cause and effect assumptions in various settings.</p>
	<p>Scalability:</p> <p>Create generalizable knowledge that facilitates the up-take of experiment results in system-wide applications</p>	<p>Scalability:</p> <p>Does the transition experiment indicate the potential for and how outputs can be scaled out to broader applications or up to higher hierarchical levels?</p>	<p>Independence of measures from changing governance systems on different scales.</p>
	<p>Accounting for unintended consequences associated with up-take:</p> <p>Reflect on and identify circumstances that have the potential to</p>	<p>Accounting for unintended consequences associated with up-take:</p> <p>Does the transition experiments account</p>	<p>Typical indicators are consideration of rebound effects, long-term consequences, and the potential for</p>

	generate unintended consequences through the up-take of sustainability solutions.	for unintended consequences that are associated with the up-take of sustainability solutions?	cooptation and offsetting of sustainability gains.
<b>Outcome</b>			
<b>Socio-ecological integrity</b>	Harmonize human well-being with the biophysical processes and elements, preventing degradation of ecosystems and reducing overall impacts and threats to the life-support system.	Do the transition experiment's outputs strengthen socio-ecological integrity?	New green walls and roofs, ecosystem-based spatial planning including adapted user behavior, and new, improved or prioritized habitat (i.e. blue and green infrastructure).
<b>Livelihood sufficiency and opportunity</b>	Ensure sufficient access of individuals and communities to what is needed for a decent life and create opportunities for positively exercising power and capabilities.	Do the transition experiment's outputs enhance livelihood sufficiency and opportunity?	Access to portable water and availability of water.
<b>Intra- and inter-generational equity</b>	Ensure sufficient and effective choices that reduce gaps between the rich and the poor and enhance opportunities of future generation to pursue sustainable lives.	Do the transition experiment's outputs improve intra- and inter-generational equity?	Creation of opportunities for various social groups, particularly those least privileged, and ensuring equity between providers and beneficiaries.
<b>Resource maintenance and efficiency</b>	Create sustainable livelihoods for all while reducing threats that jeopardize the long-term socio-ecological integrity and cutting material and energy use per unit of benefit.	Do the transition experiment's outputs contribute to overall resource maintenance and efficiency?	Cradle-to-cradle or "Benign by Design" approaches, reduction in resource consumption, and efficiency gains in agricultural energy production
<b>Socio-ecological stewardship and</b>	Provide arrangements that support individual and collective sustain-	Do the transition experiment's outputs build or support socio-ecological under-	Participatory settings, collaboration among different actors,

<b>democratic governance</b>	ability decision-making fostering ongoing collaborative actions, social inclusion and ownership.	standing and democratic governance?	knowledge coproduction, strengthened human-nature relationships, and effective public input into municipal decision making.
<b>Precaution and adaption</b>	Acknowledge uncertainty and avoid uncomprehended risks, creating learning opportunities and preparing for surprises and change.	Do the transition experiment's outputs ensure precaution and adaptation?	risk-averse and cautious approaches, comprehensive risk analysis, and measures that explicitly address environmental degradation.
<b>Process</b>			
<b>Sequence of actions</b>	Document the chronological chain of activities including the act of doing within the experiment, its purpose, the delivered actions and the scope of interventions.	Is the transition experiment structured into a meaningful sequence of actions?	Adequate planning of actions and their interference in the timeline of the experiment, the completeness of actions as well as engaging the right participants and the right information
<b>Sound methodology</b>	Ensure that the experiment is facilitated through sound methods, including problem analysis, visioning, strategy development, as well as monitoring and evaluation	Does the transition experiment adopt a sound methodology to conduct the experiment?	Structured procedures for generating outputs and the adequacy of methods for the respective action.
<b>collaboration</b>	Provide participatory settings for collaboration of participants and ensure empowerment of participants.	Does the transition experiment facilitate collaboration among relevant stakeholders in the experimentation process?	Affiliations of participants and their roles, information flows, decision-making procedures, and interactions.
<b>Reflexivity and learning</b>	Ensure the analysis of actions, structures, processes and out-	Does the transition experiment foster reflexivity and learning	Presence of a shared learning agenda and dedicated points of reflections such as meetings to explicitly

	puts, as well as iterative and recursive learning.	throughout the process?	reflect on the experiment, review processes, as well as changes of the experimentation process.
<b>transparency</b>	Ensure open and truthful reporting on intentions and pursued actions within the experimentation process.	Does the transition experiment ensure transparency throughout the process?	Openly published results, reports that explicate assumptions and intentions, and documentation of the decision-making process.
<b>Input</b>			
<b>Awareness</b>	Enable participants' consciousness of and ability to acknowledge the need for radical real-world changes prior to their engagement in the experiment.	Does the transition experiment involve participants that are aware of the need for transformational change pursued through the experiment?	Sustainability related track records of participants, and participants' general awareness of the sustainability issues tackled by the experiment.
<b>Commitment</b>	Cater for willingness, promises, positive attitudes and interests of involved participants to explore intentionally radical instead of incremental changes	Does the transition experiment involve participants committed to carrying out the experiment?	Participants' agreement to deliver tasks on time, participants' engagement in decision-taking, and continuous participation in the experimentation.
<b>Expertise</b>	Ensure expertise of participants in sustainability transition experiments including widely recognized professional skills and experiential techniques to research, craft, guide, decide and judge experimentation.	Does the transition experiment involve participants who possess the necessary skills and knowledge to carry out the experiment?	Related work experience and academic and professional degrees and training of the participants.
<b>Trust</b>	Cater for mutual willingness of and between researchers and non-academic	Does the transition experiment involve	Participants' attitudes toward other participants, ability to speak

	participants to rely on actions of other members of the sustainability transition experiment.	participants who trust each other?	one's mind, and willingness to rely on others' judgments and capacities.
<b>Support</b>	Ensure structural, financial and nonfinancial resources as well as assistance from public and private authorities in preparing and executing sustainability transition experiments.	Does the transition experiment secure sufficient support for the experimentation?	Available funds, positions, hours of voluntary contributions and endorsements from actors and institutions.

**Tab. 11: Gütekriterien für Methoden der partizipativen Wissenserzeugung in der transdisziplinären Forschung.** Die dargestellten Gütekriterien sind hier nach Ziel und Kontext, Umsetzung der Methoden und Erfahrung im Einsatz der Methoden geordnet. Die ursprünglichen Erläuterungen der Autor\*innen wurden in dieser Darstellung als Fragen umformuliert (eigene Darstellung nach Defila und Di Giulio 2018a, S. 55ff.).

Gütekriterien	Weiterführende Fragen/Erläuterungen
<b>Ziel und Kontext</b>	
<b>Zeitpunkt im Forschungsdesign/Projektlauf – wann?</b>	<p>Wann lässt sich die Methode im Projektverlauf besonders gut einsetzen? (relativ zu den Arbeiten und zum Projekt, das heißt in welchem Projektschritt, nicht als absoluter Zeitpunkt)</p> <p>Wann lässt sich die Methode gar nicht einsetzen?</p>
<b>Angestrebtes Ziel – Wozu?</b>	<p>Welchem Ziel dient die Methode?</p> <p>Welche Frage soll damit beantwortet werden?</p> <p>Welche Art von Ergebnis kann damit erzeugt werden?</p> <p>Welche Art von Wissen kann damit produziert werden? (Systemwissen/ Zielwissen/ Transformationswissen)</p> <p>Bei einer Methode der Konsens- oder Synthesebildung: Welche Art von integriertem Wissen kann damit erzeugt werden? (z.B.: mehrperspektivische Analyse eines Phänomens, Meta-Theorie, integriertes Begriffssystem, gesamtheitliche Systemdarstellung)</p>
<b>Partizipationsgrundlage – wie</b>	Welche Voraussetzungen der Partizipation sollten im Praxisfeld bzw. bei den partizipierenden Praxisakteuren bestehen, damit die Methode gut und zielführend umgesetzt werden kann?

<b>sieht das Feld aus?</b>	<p>Welche Rahmenbedingungen wären/sind im Praxisfeld besonders förderlich bzw. hemmend?</p> <p>Welche Ressourcen, Kompetenzen und Bereitschaften sind seitens der Praxisakteure für die Umsetzung der Methode nötig?</p> <p>Wie zugänglich und verständlich ist die Methode für die Praxisakteure?</p> <p>Welche Art von Akteuren soll mit der Methode erreicht werden? Wer könnte davon ggf. abgeschreckt werden?</p> <p>Wie kann mit Qualifikations- und Funktionsunterschieden umgegangen werden?</p> <p>Wie kann mit Konflikten im Praxisfeld umgegangen werden?</p> <p>Sind Unterstützung/Schulungen für die Praxisakteure ggf. notwendig und wenn ja, welche?</p>
<b>Umsetzung</b>	
<b>Ablauf und Tätigkeiten – wie?</b>	<p>Wie wird die Methode durchgeführt? Welche Handlungen müssen in welcher Reihenfolge stattfinden?</p> <p>Welche Voraussetzungen müssen gegeben sein/geschaffen werden, um die jeweiligen Handlungen auszuführen? (in Hinblick auf konkrete Techniken, Infrastrukturen, Software, Räumlichkeiten, Material)</p> <p>Welche möglichen Gefahren und Stolperfallen gibt es ggf.?</p> <p>Wie sollen Daten aus den Zwischenschritten verarbeitet und weiterverwendet werden?</p> <p>Gibt es weitere Varianten des Vorgehens? Wenn ja, wonach wird entschieden welche Variante gewählt wird bzw. was muss berücksichtigt werden?</p>
<b>Ethik – was darf man?</b>	<p>Was ist aus ethischer Sicht relevant in Hinblick auf die Umsetzung der Methode und die anschließende Aufbereitung und Weiterverwendung der gewonnen Daten/Ergebnisse?</p> <p>Welche Vorkehrungen müssen zum Schutz, der direkt oder indirekt an der Methodenumsetzung bzw. von der Methodenumsetzung betroffenen Personen, beachtet werden?</p> <p>Welche Handlungen sind ethisch bedenklich/nicht zulässig?</p> <p>Welche Prinzipien, Regeln und Kriterien müssen ggf. bei welchem Schritt beachtet werden?</p> <p>Wo sind ggf. Grenzen zu ziehen bei der Umsetzung der Methode, damit die Praxisakteure keinen Sachzwängen unterworfen sind, die sie im Vorfeld nicht abschätzen konnten?</p>

<b>(Vorwissen- schaftlicher) Ge- genstand – was wird erfasst?</b>	<p>Welche (vorwissenschaftlichen) Gegenstände soll mit der Methode in den Blick genommen werden?</p> <p>Anhand welcher Kriterien kann geprüft werden, ob die Methode dafür adäquat ist?</p> <p>Welche Art von Daten lassen sich daraus erzeugen? Ergeben sich aus der Art der erzeugten Daten Konsequenzen für die Verarbeitung und Weiterverwendung?</p> <p>Auf welche Art von vorwissenschaftlichen Gegenständen sollte die Methode auf keinen Fall angewendet werden bzw. welche Aspekte von Gegenständen sollten ausgeschlossen werden, da sich die Methode hierfür nicht eignet?</p>
<b>Zugrunde lie- gende Vorstel- lungen – wel- ches Weltbild, welches Men- schenbild, wel- che Theorie?</b>	<p>Welche theoretischen Positionen und Diskurse liegen der Methode zu- grunde?</p> <p>Welche Vorbehalte könnten diesbezüglich ggf. bestehen?</p> <p>Welche anderen Theorien lassen sich damit evtl. gut in Verbindung setzen, welche nicht?</p>
<b>Kombination – wie sind die Me- thoden kombi- niert?</b>	<p>Falls es sich um eine Einzelmethode handelt: Gibt es (Arten von) Methoden mit der sich die Methode gut kombinieren lässt? Gibt es (Arten von) Methoden, die ggf. mit der Methode assoziiert werden können, aber nicht kombiniert werden sollten?</p> <p>Falls es sich um eine, aus anderen Methoden rekombinierte Methode handelt: Welche Einzel-Methoden(-Teile) wurden kombiniert? Was macht die Konsistenz der Methoden-Kombination aus, in Hinblick auf Ziel, Gegenstand, Theorien und soziale Interaktion? Worauf ist bei der Umsetzung zu achten, damit kohärente Ergebnisse und keine vielfältigen, unzusammenhängenden Teilergebnisse entstehen?</p>
<b>Beziehung zwi- schen Forsch- enden und Praxis- akteuren – wie Nähe und Dis- tanz gestalten?</b>	<p>Wem kommt welche Rolle zu bei der Umsetzung der Methode?</p> <p>Welche Entscheidungen/Handlungen sollten von Forschenden und welche von Praxisakteuren getroffen werden und warum?</p> <p>Welche Beziehung wird zwischen den Forschenden und Praxisakteuren vorausgesetzt und wie wird sie ggf. durch die Methode verändert?</p> <p>Wo besteht die Gefahr, dass die nötige Distanz zwischen Forschenden und Praxisakteuren verloren geht und wie lässt sich in diesem Fall intervenieren?</p> <p>In welchem Fall ist darauf zu achten, die Unterschiede zwischen Forschenden und Praxisakteuren nicht zu verwischen?</p> <p>Wo besteht während der Umsetzung Reflexions-Bedarf?</p>

	Welche Instrumentalisierungsrisiken sind zu vermeiden?
<b>Beziehung der Praxisakteure untereinander – wie mit Macht und Ausschluss umgehen?</b>	<p>Welche Beziehung wird zwischen den Praxisakteuren sowie den Praxisakteuren und dem Praxisfeld wird vorausgesetzt und wie wird sie ggf. durch die Methode verändert?</p> <p>Welche Beziehungen könnten die Umsetzung der Methode gefährden?</p> <p>Welche Machtstrukturen innerhalb und außerhalb des Projekts könnten sich verzerrend auf die Umsetzung der Methode auswirken oder ihr gar entgegenwirken?</p> <p>Welche Formen des Ausschlusses sind für die Methode relevant und wie ist diesen zu begegnen?</p> <p>Wo könnte in der Umsetzung der Methode die Gefahr aufkommen, dass sich die Praxisakteure nicht ausreichend von ihrem Praxisfeld und ihrer Rolle darin distanzieren? Wie könnte dieses Risiko verringert werden?</p> <p>Wo besteht während der Umsetzung Reflexions-Bedarf?</p>
<b>Dokumentation – wie Transparenz herstellen?</b>	<p>Welche Entscheidungen bei der Vorbereitung, Umsetzung oder Nachbereitung der Methode müssen dokumentiert werden?</p> <p>Wer muss über welche Entscheidungen informiert werden?</p> <p>Welche Daten und (Zwischen-)Ergebnisse müssen dokumentiert werden?</p> <p>Welche Dokumentation muss den Forschenden und/oder Praxisakteuren zur Verfügung gestellt werden? (in Hinblick auf den Schutz der Beteiligten und Schutz geistigen Eigentums)</p> <p>Wie soll die Dokumentation erfolgen?</p> <p>Wie soll über das Vorgehen und die Ergebnisse nach Außen berichtet werden?</p>
<b>Erfahrungen im Einsatz der Methoden</b>	
<b>Aufnahme der Methode – wie kommt sie an?</b>	<p>Was wurde als positiv erachtet, ist bei der Umsetzung besonders leichtgefallen?</p> <p>Was wurde als negativ erachtet, hat bei der Umsetzung eher Mühe bereitet?</p> <p>Welche Konsequenzen ergeben sich daraus für die Wahl und/oder die Anpassung und Umsetzung der Methode?</p> <p>Welche Instrumente könnten verwendet werden/auf welche Indizien sollte geachtet werden, um in der Umsetzung der Methode feststellen zu können, wie die Praxisakteure auf die Methode reagieren?</p>

---

<b>Angemessenheit – welche Erwartungen, welche möglichen Schäden?</b>	Welche Erwartungen wurden durch die Methode geweckt?
	Welchen Erwartungen entspricht die Methode bzw. welche können weniger/nicht eingeschlossen werden durch die Methode?
	Welche Schäden können durch die Methode verursacht werden? An welchen Kriterien wird dies gemessen? Wie können die Schäden vermieden werden?
	Gibt es Praxisfelder oder Konstellationen von Praxisakteuren, die eine Umsetzung der Methode verunmöglichen oder nicht ratsam sind?

---

[www.plastik-reduzieren.de](http://www.plastik-reduzieren.de)

