

Berücksichtigung der Leistungen für Praxis und Gesellschaft in der Forschungsevaluierung

Evaluating the impact of research on practice and society

FKZ: 06OE307

Projektnehmer:

Universität Kassel
FB 11, Fachgebiet Ökologischer Land- und Pflanzenbau
Nordbahnhofstraße 1a, 37213 Witzenhausen
Tel.: +49 5542 98-1565
Fax: +49 5542 98-1568
E-Mail: birge.wolf@uni-kassel.de
Internet: <http://www.uni-kassel.de>

Autoren:

Wolf, Birge; Lindenthal, Thomas; Szerencsits, Manfred; Heß, Jürgen

Gefördert vom Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz
im Rahmen des Bundesprogramms Ökologischer Landbau und andere Formen nachhaltiger
Landwirtschaft (BÖLN)

Berücksichtigung der Leistungen für Praxis und Gesellschaft in der Forschungsevaluierung

„Praxis-Impact“ Synthesebericht

06 OE 307

Birge Wolf
Thomas Lindenthal
Manfred Szerencsits
Jürgen Heß

Fachgebiet
Ökologischer Land- und Pflanzenbau
Nordbahnhofstr. 1a
37213 Witzenhausen



Ökologische Agrarwissenschaften **U N I K A S S E L**

Gefördert durch die BLE, im Rahmen des

BÖLN

Bundesprogramm Ökologischer Landbau
und andere Formen nachhaltiger
Landwirtschaft

Kurzfassung

Die Leistungen der (Ökolandbau-) Forschung für Praxis und Gesellschaft in der Forschungsevaluierung berücksichtigen: Kriterien und ein Dokumentationskonzept

Birge Wolf, Thomas Lindenthal, Manfred Szerencsits, Jürgen Heß
Fachgebiet Ökologischer Land- und Pflanzenbau, FB 11 Universität Kassel Witzenhausen, Nordbahnhofstr. 1a, 37213 Witzenhausen
birge.wolf@uni-kassel.de

Im etablierten System der Forschungsevaluierung (z. B. Peer Review Publikationen, SCI, SSCI, JIF) werden die Wirkungen der Wissenschaft auf Praxis und Gesellschaft nur unzureichend berücksichtigt. Im BÖLN Projekt 06 OE 307 wurde daher erforscht, wie praxis- und gesellschaftsrelevante Leistungen der Ökolandbauforschung adäquat in der Evaluierung berücksichtigt werden können. Die Analyse von Konzepten für die Evaluierung angewandter, insbesondere inter- und transdisziplinärer Forschung und qualitative Interviews mit 22 Experten der ökologischen Agrarforschung zeigten, mit großer Übereinstimmung, dass folgende Leistungen der Forschung für eine solche Evaluierung relevant sind:

- Fragestellungen und umsetzbare Ergebnisse haben einen Bezug zu relevanten Problemlagen von Praxis und Gesellschaft;
- Zusammenarbeit und Austausch mit anderen Wissenschaftlern und Akteuren aus Praxis und Gesellschaft;
- Publikationen und andere Produkte der Forschung für nicht-wissenschaftliche Zielgruppen,
- Relevanz und tatsächlicher Impact für Praxis und Gesellschaft.

Von den Interviewpartnern wurde ein breites und kontextbezogen anpassbares Kriterien-set für notwendig erachtet. Dieses stimmt mit Empfehlungen in der Evaluierungsliteratur überein. Die meisten bestehenden Konzepte sind spezifisch auf den jeweiligen Evaluierungsgegenstand abgestimmt und unterscheiden sich daher in Zielsetzungen, Methoden sowie Breite und Detaillierungsgrad der Kriterien-sets und Bewertungsraster. Entsprechend braucht eine generelle Ergänzung der Forschungsevaluierung anpassbare Kriterien. Die Konzepte in der Literatur umfassen vielfältige Evaluierungsinstrumente, u.a. Multi-Methoden Ansätze und die Kombination interner und externer Perspektiven. Diese Evaluierungskonzepte werden meistens als „stand-alone-Vorhaben“ umgesetzt. Dadurch erfolgt die Datenerhebung derzeit mit beträchtlichem Aufwand (z.B. Interviews oder Dokumentenanalysen). Erste Ansätze einer generalisierten Datenerfassung finden sich beispielsweise bei den Research Councils in Großbritannien. Von den projektbeteiligten Agrarwissenschaftlern wurden drei Aspekte für die Methodik der Evaluierung besonders betont:

- a) Akteure aus Praxis und Gesellschaft sollten an der Evaluierung beteiligt sein;
- b) Für die Analyse von Folgewirkungen ist eine angemessene Nachbeobachtung über das Projektende hinaus sinnvoll;
- c) Zur Verbesserung der Anwendbarkeit und Akzeptanz der Evaluierung sollte der zusätzliche Aufwand von Forschern für die Datenerfassung gering bleiben und die Daten für verschiedene Evaluierungen nutzbar gemacht werden.

Im vorliegenden Projekt wurde ein Konzept entwickelt, mit dem die Datenerfassung für die Evaluierung in die Antragstellung und Berichterstattung für die öffentliche Drittmittelförderung integriert werden könnte. Dieses würde sowohl einen zusätzlichen Aufwand für die Datenerfassung seitens

der Wissenschaftler und der Evaluatoren vermeiden, als auch die Nutzbarkeit der Daten verbessern. Ein solches Vorgehen ist naheliegend, weil eine Analyse sowohl der Leitbilder als auch der Dokumentationsanforderungen vieler öffentlicher Forschungsförderer (z.B. BMBF, BMWi, BMELV, BMU) zeigte, dass eine erhebliche Schnittmenge mit dem Datenbedarf für die Evaluierung von Leistungen der Forschung für Praxis und Gesellschaft gegeben ist. Um die bereits vorhandenen Informationen in Anträgen und Berichten für eine Evaluierung nutzbar zu machen, ist es notwendig,

- a) die Daten so zu strukturieren, dass sie gleichermaßen für die Forschungsförderung und verschiedene Evaluierungsgegenstände (Wissenschaftler, Projekte, Institutionen, Programme) nutzbar sind,
- b) bisher nicht oder nur implizit geforderte, aber für die Evaluierung notwendige Informationen, explizit zu erfassen und
- c) für die technische Umsetzung ein Datenbanksystem zu nutzen.

Durch letzteres können einmalig eingegebene Daten leicht zu späteren Zeitpunkten ergänzt werden (Nachbeobachtung) und Datensätze kontextbezogen für verschiedene Evaluierungsgegenstände gefiltert und unterschiedlich detailliert oder aggregiert ausgewertet werden. Da Teile von Anträgen und Berichten durch strukturierte Dokumentation ersetzt werden können, bliebe der Dokumentationsaufwand für Wissenschaftler voraussichtlich gleich. Wie und mit welchem Nutzen das skizzierte Konzept umgesetzt werden kann, wird über Fallstudien in einem Folgeprojekt erprobt. Darüber hinaus ist eine breite Diskussion in der Scientific Community weiterhin notwendig um eine Ergänzung der Forschungsevaluierung über den wissenschaftlichen Impact hinaus zu gestalten.

Abstract

Evaluating the impact of research on practice and society. Contributions for a complementary assessment from the perspective of organic agriculture.

Birge Wolf, Thomas Lindenthal, Manfred Szerencsits, Jürgen Heß
Organic Farming and Cropping Systems, Organic Agricultural Sciences, University Kassel,
Nordbahnhofstr. 1a, 37213 Witzenhausen
birge.wolf@uni-kassel.de

The established system of research evaluation (e.g. SCI) does not sufficiently consider the impact of science on practice and society. Therefore the BÖLN-Project (06OE307) focused on the adequate extension of research evaluation beyond scientific impact. The project was carried out exemplarily for organic agricultural science.

Published evaluation concepts for applied, interdisciplinary and transdisciplinary research and interviews with 22 agricultural scientists shows a high level of agreement that the following achievements of research are relevant for an evaluation of the impact on practice and society:

- Research questions and applicable results are related to relevant problems of practice and society;
- Cooperation with scientists particularly from other disciplines, cooperation and knowledge transfer with actors from practice and society;
- Publications and products of research for non-scientific target groups;
- Relevance and the subsequent impact on practice and society.

Interviewees propose a comprehensive and adaptable criteria set for the evaluation of the impact on practice and society. This corresponds to recommendations in the evaluation literature. Most concepts are tailored for specific evaluation objects and purposes. Accordingly, they differ considerably in focus and extent of criteria sets. Thus, general extension needs adaptability of criteria sets.

Literature concepts also provide a broad spectrum of evaluation methods. Frequently multi-method-approaches are used, which often consider both internal and external perspectives. In most cases evaluations were carried out as “stand-alone-procedures” with high effort for data collection, like interviews or evaluation of project reports. Attempts for a regular data assessment are made for example by the research Councils UK. Indeed, project results about methods point out that:

- a) Actors from practice and society have to participate in the evaluation process.
- b) A consequent follow-up-assessment for projects is needed to gather later impact on practice and society.
- c) Data should be collected in a less time-consuming manner and be available for various evaluations and other purposes to enhance usability and acceptance of complementary assessment procedures.

To improve data collection and usability for evaluation (issue c), we developed a concept to integrate data collection in the application and reporting of third-party funded research projects. We did so, since we found considerable overlapping between data requirements for the evaluation of the im-

pact on practice and society and the documentation requirements and guidelines of many public research funding organisations (e.g. BMBF, BMWi, BMELV, BMU).

To enhance the usability of information provided in proposals and reports for evaluation, it is necessary to

- a) structure the data in such a way that makes information useable both for research funding as well as for evaluations,
- b) explicitly gather information necessary for the evaluation which is currently not or only implicitly claimed and
- c) store it in a database-system.

The latter enables the use of data for various evaluation and research administration purposes, because it can be filtered, aggregated and analysed as needed. Furthermore it can easily be updated with information from follow-up-assessments.

If structured documentation by scientists may substitute considerable parts of research proposals and reports, overall efforts of scientists for documentation would not increase. To further develop this concept in detail it will be implemented and tested in case studies in a follow-up project. Practical implementation would also allow verifying, if the intended benefits for scientists, research funding organisations and evaluation can be achieved. Additionally, a wide discussion within the Scientific Community is needed, to shape an general extension of research evaluation beyond scientific impact.

Vorbemerkung

Dieser Synthesebericht spiegelt den Stand des Wissens und die Ergebnisse zum Zeitpunkt der Fertigstellung des Berichts im September 2011 wieder.

Die weiterentwickelten Ergebnisse dieses Projektes gingen in eine referierte Publikation ein (open access).

Wolf, B. , Szerencsits, M. , Lindental, T., Hoolbrook, J.B. , Heß, J. (2013): Research Evaluation beyond Scientific Impact. How to Include Criteria for Productive Interactions and Impact on Practice and Society. GAIA 22/2, S.104–114. <http://www.ingentaconnect.com/content/oekom/gaia/2013/00000022/00000002/art00009>

Desweiteren wurde auf Basis der Ergebnisse ein Folgeprojekt initiiert.

Titel: Weiterentwicklung und Erprobung eines Konzeptes zur Dokumentation und Evaluierung von Leistungen der Agrarforschung für Praxis und Gesellschaft. (Praxis-Impact II), 2012 – 2015. Gefördert durch das BMELV im Rahmen des Bundesprogramms Ökologischer Landbau und andere Formen nachhaltiger Landwirtschaft.

<http://www.uni-kassel.de/fb11agrар/fachgebiete-einrichtungen/oekologischer-land-und-pflanzenbau/forschung/laufende-projekte/praxis-impact-ii-2012-2015.html>

Im Text wird mehrfach auf verschiedene Anhänge verwiesen, die jedoch nicht veröffentlicht werden. Diese Anhänge haben dokumentarischen Charakter und dienen als Arbeitsgrundlage für ein Folgeprojekt. Auf Anfrage können einzelne Anhänge bei den Autoren zur internen Nutzung angefordert werden. Die Anhänge umfassen insgesamt:

- Anhang I:** kritisches Review der bestehenden Forschungsevaluierung
- Anhang II:** Detaillierte Auswertung der Interviews und Workshops
- Anhang III:** Datenset zur strukturierten Dokumentation von Leistungen der Forschung für Praxis und Gesellschaft
- Anhang IV:** Dokumentationsleistungen für die BLE (u.a. Protokolle der Workshops, Vortragsunterlagen, Publikationen)

Nicht zuletzt wollen wir darauf verweisen, dass im Text der Einfachheit halber nur die männliche Form verwendet wird, die weibliche Form aber stets mit eingeschlossen ist.

Danksagung

Ein ganz besonderer Dank gilt den Interviewpartnern und Workshopteilnehmern des Projektes! Sie haben ihre umfangreichen Kenntnisse und Erfahrungen an uns weiter gegeben – und waren bereit die dafür notwendige Zeit zu investieren. Einige sind, über die Interviews und Workshops hinaus, zu kontinuierlichen Diskussionspartnern geworden und haben das Projekt und seine Zielsetzung über die gesamte Laufzeit unterstützt.

Besonders hervorheben möchten wir die engagierte Mitarbeit von Prof. Hermann Boland, insbesondere die methodische Begleitung des Projektes. Ebenfalls persönlich bedanken möchten wir uns für die Initiative von Dr. Thorsten Michaelis, Prof. Werner Zollitsch und PD Dr. Stephan Albrecht.

Durch die Teilnehmer der Transfer-Workshops und Vorträge wurde unser Vorhaben durch vielfältige, projekt-externe Perspektiven ergänzt. Wir sind sehr glücklich darüber, dass sie auf diesem Wege entschieden zur Präzisierung der Ergebnisse und Weiterentwicklung des Konzeptes beigetragen haben. Vielen Dank. Besonderer Dank gilt dabei den jeweiligen Leitern der Institutionen, die diese Transfer-Workshops ermöglicht und an den jeweiligen Standorten vorbereitet haben.

Dem Bundesprogramm Ökologischer Landbau und andere Formen nachhaltiger Landwirtschaft danken wir für die Finanzierung des Projektes. Besonders anerkennen möchten wir auch die Arbeit der Projektadministration, die eine flexible Anpassung des Projektverlaufs stets ermöglicht hat.

Inhalt

1	Einleitung Problemstellung und Zielsetzung des Projektes	1
1.1	Einleitung	1
1.2	Problemstellung.....	1
1.3	Zielsetzung	3
1.4	Begriffsklärung „Praxis“, „Gesellschaft“, „praxisorientiert“	4
2	Projektablauf, Material und Methoden.....	5
2.1	Projektphase 1 - Entwicklung eines Leitfadens für die Evaluierung praxisbezogener Ökolandbauforschung	5
2.2	Projektphase 2 - Transfer und Diskussion von Projektergebnissen.....	6
2.3	Projektphase 3 - Konzept für die strukturierte Dokumentation von Leistungen der Forschung für Praxis und Gesellschaft in Anträgen und Berichten	6
3	Forschungsevaluierung: Einführung und Kritik	8
3.1	Einführung in die Evaluierung	8
3.1.1	Begriffe und Grundlagen der Evaluierung	8
3.1.2	Standards für die Evaluation	10
3.1.3	Erkenntnistheoretische Fundamente der Evaluation	11
3.1.4	Diskussion des eigenen Evaluierungsparadigmas	11
3.2	Einführung in die Forschungsevaluierung.....	12
3.3	Das etablierte Evaluierungssystem in der kritischen Diskussion.....	14
3.3.1	Kritikpunkte am etabliertem Evaluierungssystem in der Literatur.....	14
3.3.2	Einschätzung des etablierten Evaluierungssystem in den Interviewergebnissen.....	17
4	Berücksichtigung von Leistungen der Forschung für Praxis und Gesellschaft in der Forschungsevaluierung.....	19
4.1	Stand des Wissens zur Berücksichtigung von Leistungen für Praxis und Gesellschaft in der Forschungsförderung und Forschungsevaluierung.....	19
4.1.1	Ziele öffentlich geförderter Forschungsprogramme der Agrar- und Nachhaltigkeitsforschung	19
4.1.2	Anforderungen im Zuge der Beantragung von Drittmitteln in der EU und USA.....	19
4.1.3	Ausgewählte Leitbilder der deutschen Bundesministerien.....	19
4.1.4	Evaluierungskonzepte für die inter- und transdisziplinäre Forschung	21
4.2	Ergebnisse der Befragung und Gruppendiskussion diskutiert mit einschlägiger Literatur. ..	22
4.3	Anforderungen und Ziele für die Evaluierung praxisorientierter Ökolandbauforschung.....	23
4.3.1	Individuelle Anpassung der Evaluierung	23
4.3.2	Bewertung verschiedener Kriterienbereiche	24
4.4	Kriterienvorschläge in Interviews, Workshop und Literatur.....	27
4.5	Vorgeschlagene Evaluierungsinstrumente im Kontext der Literatur	34

4.5.1	Spektrum der Methoden für die Bewertungsgrundlagen und den Bewertungsprozess .	34
4.5.2	Vertiefung: Praxiseinbindung.....	37
4.6	Zwischenfazit aus Projektphase I und II - Identifikation von Möglichkeiten der breiteren Etablierung	41
5	Konzeptentwicklung für die strukturierte Dokumentation von Leistungen für Praxis und Gesellschaft in Anträgen und Berichten der öffentlichen Forschungsförderung	42
5.1	Die vorhandene Dokumentation von Forschungsleistungen für die Evaluierung	42
5.2	Integration der Datenerfassung in die Antragsstellung und Berichterstattung.	44
5.3	Konzept für die strukturierte, datenbankgestützte Dokumentation in Anträgen und Berichten	47
5.3.1	Das vorgestellte Konzept im Vergleich zur etablierten Forschungsevaluierung.....	51
6	Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse.....	53
7	Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen	54
8	Zusammenfassende Schlussfolgerungen und Perspektiven.....	55
9	Literaturverzeichnis.....	57
10	Wissenstransfer	63
10.1	Erfolgte und geplante Veröffentlichungen	63
10.2	Aktivitäten zur Verbreitung der Ergebnisse	63

Abkürzungsverzeichnis

WS	Workshopteilnehmer am Ergebnisworkshop 01/2010
Q	Quantitative Frage im Interviewleitfaden

Abbildungsverzeichnis

Abbildung 1: Gesellschaftlicher Nutzen in Leitbildern öffentlich geförderter Forschung, in der etablierten Forschungsevaluierung und die Anreize für die Ausrichtung der Forschung.	2
Abbildung 2: Langfristiges Ziel des Projektes: Forschungsevaluierung berücksichtigt den „Scientific Impact“ und den „Impact on Practice and Society“, mit entsprechenden Lenkungseffekten auf die Ausrichtung der Forschung.	3
Abbildung 3: Individuelle Anpassung der Evaluierung	24
Abbildung 4: Übersicht über die Evaluierungskriterien und die Verfügbarkeit der für die Bewertung benötigten Informationen in verschiedenen Phasen des „Projektzyklus“	33
Abbildung 5: Wichtige Eckpunkte für eine gelingende Einbindung von Akteuren aus Praxis und Gesellschaft in die Forschungsevaluierung.	37
Abbildung 6: Die bisherige Forschungsdokumentation und ihre Nutzung für Forschungsevaluierung und Drittmittelakquise.	43
Abbildung 7: Leistungen für Praxis und Gesellschaft werden bereits jetzt in Anträgen und Berichten dokumentiert und könnten auch für die Evaluierung genutzt werden.	47
Abbildung 8: Nutzung der strukturierten Dokumentation für die Evaluierung verschiedener Evaluierungsgegenstände.	49
Abbildung 9: Beispiel: Vereinfachte Berichterstattung und Projekt-Administration durch strukturierte Dokumentation.	50
Abbildung 10: Evaluierung des wissenschaftlichen Impacts in einem wissenschaftsinternen, “geschlossenen” System. Manuskript > Peer Review > Publikation > Zitation als linearer Prozess (links). Ebbvaluierung der Leistungen für Praxis und Gesellschaft in einem „offenen System“ mit wechselseitigen Interaktionen und außerwissenschaftlichen Ursachen für Impact.	53

Tabellenverzeichnis

Tabelle 1: Glossar zu den wichtigsten zitationsbasierten Indikatoren der etablierten Forschungsevaluierung.....	13
Tabelle 2: Zusammenfassung der wichtigsten Kritik am etablierten Evaluierungssystem aus Sicht der Interviewpartner des Projektes.	18
Tabelle 3: Zielsetzungen ausgewählter öffentlicher Förderprogramme im Bereich Agrar- und Nachhaltigkeitsforschung.....	20
Tabelle 4: Ausgewertete Literatur zur Evaluierung inter- und transdisziplinärer (Nachhaltigkeits-)forschung sowie praxisorientierter Agrarforschung.....	21
Tabelle 5: Kriterienbereiche aus dem Interviewleitfaden - Bewertung und wichtige inhaltliche Aussagen	26
Tabelle 6: Kriterien für die Bewertung praxisorientierter Leistungen von Wissenschaftlern, Projekten, Institutionen und Programmen.	28
Tabelle 7: Kriterien für die Bewertung der Interdisziplinarität.....	29
Tabelle 8: Kriterien für die Bewertung von Nachhaltigkeit, gesellschaftlicher Relevanz, Systemorientierung, Vielfalt von Themen und Methoden, Innovation und Originalität und Langfristigkeit der Forschung.	31
Tabelle 9: Bewertungsgrundlagen und Bewertungsprozesse für die Evaluierung – Vorschläge aus den Projektergebnissen (x) und der Literatur (Zahlen aus Tabelle 4).....	35
Tabelle 10: Anforderungen für die Berichterstattung und Beantragung öffentlicher Forschungsgelder des BMBF, BMWi, BMFSFJ, BMELV, BMU, BLE (BLE, BMBF 2006) (links) und für den Datenbedarf zur Evaluierung wissenschaftlicher Leistungen für Praxis und Gesellschaft (rechts).	45
Tabelle 11: Ergänzungs- und Anpassungsbedarf, um die bisherige Dokumentation in Anträgen und Berichten für die Forschungsevaluierung nutzbar zu machen.....	46
Tabelle 12: Vergleich der Evaluierung des „Scientific Impact“ und dem vorgestellten Konzept für die Evaluierung der Wirkungen in Praxis und Gesellschaft.....	52

1 Einleitung Problemstellung und Zielsetzung des Projektes

1.1 Einleitung

Derzeit wird international und fachübergreifend vielfältig diskutiert, wie die Forschung verstärkt Beiträge zur Lösung gesellschaftlicher Probleme und nachhaltigen Entwicklung leisten kann. Der (öffentlich geförderten) Wissenschaft wird dabei sowohl die Verantwortung gegenüber der Gesellschaft als auch die Fähigkeit zugesprochen, hierfür Beiträge zu leisten. Gleichzeitig werden aber auch Strukturen, Entwicklungen oder Förderschwerpunkte genannt, die einer nachhaltigen Entwicklung abträglich sind. Vielfach werden fundamentale Veränderungen in der Forschungsförderung, wie auch im Selbstverständnis der Forschung gefordert, damit die Forschung ihrer Verantwortung gegenüber der Gesellschaft gerecht wird und ihre Potenziale zur Lösung gesellschaftlicher Probleme und einer nachhaltigen Entwicklung verstärkt ausschöpfen kann. (Vgl. „science for global sustainability“ Clark, Crutzen und Schellnhuber 2005, VDW 2010b, WBGU 2011, 341–83, VisionRD4SD 2011).

In Bezug auf die Agrarwissenschaften werden beispielsweise gleichberechtigte Partnerschaften zwischen Wissenschaft und Praxis als Voraussetzung gesehen, damit die Forschung zur weltweiten Ernährungssicherung auf ökologisch und sozial nachhaltige Weise beiträgt (IAASTD, 2009, 20, 33). In der Forschung zur Ökologischen Landwirtschaft wird die Partizipation der Praxis sogar als wichtiges Merkmal und Zukunftsperspektive für ihren Erfolg gesehen (Niggli and Gerber 2010).

Einige Reflektionen zur Förderung und zum Selbstverständnis der Forschung umfassen auch eine veränderte Forschungsevaluierung, um Anreize für die Forschung zu setzen, verstärkt zu einer Lösung gesellschaftlicher Probleme beizutragen. Des Weiteren werden notwendige Veränderungen auch aus der Kritik am etablierten Evaluierungssystem heraus formuliert.

Gerade im Agrar- und Ernährungssektor mit seiner existenziellen Bedeutung für das menschliche Leben, muss die Forschung einen Beitrag zur Lösung von Zukunftsproblemen leisten. Eine veränderte Forschungsevaluierung wird daher im Weltagrarbericht explizit für die Agrarforschung gefordert und zwar im Hinblick auf eine Berücksichtigung von Gerechtigkeit, ökologischer und sozialer Nachhaltigkeit (IAASTD 2009, 217).

1.2 Problemstellung

In der Literatur wird vielschichtige Kritik am etablierten Evaluierungssystem dokumentiert (siehe Kap. 3.3), die von den Interviewpartnern des Projektes teilweise auch in Bezug zur Forschung für den Ökologischen Landbau als relevant eingeschätzt wurde. Unabhängig davon, inwieweit diese Kritik von einzelnen Vertretern der Scientific Community geteilt wird, kann die Ausgangslage folgendermaßen beschrieben werden (siehe auch Abbildung 1):

- Viele Leitbilder öffentlicher Forschung formulieren als Zielsetzung sowohl wissenschaftliche Qualität als auch den gesellschaftlichen Nutzen in vielfacher inhaltlicher Ausgestaltung (vgl. Kap. 4.1).
- Die wissenschaftliche Qualität wird im etablierten Evaluierungssystem durch das Peer Review und standardisierte quantitative Indikatoren erfasst und bewertet. Da jede Form der Bewer-

tung Lenkungseffekte für die Erbringung der bewerteten Leistungen hat, fördert das derzeitige System die Konzentration auf die gemessenen wissenschaftlichen Leistungen.

- Da der gesellschaftliche Nutzen der Forschung in der etablierten Forschungsevaluierung bisher nur unzureichend berücksichtigt wird, werden zum einen nicht genügend Anreize gesetzt, die auf eine Erbringung eines gesellschaftlichen Nutzens gezielt hinwirken. Zum anderen bleibt der gesellschaftliche Nutzen der Forschung dadurch eine weitgehend unbekannte Größe.
- Welchen Nutzen die Wissenschaft für die Gesellschaft hat, wird vielschichtig diskutiert (vgl. Schavan 2008). Einerseits wird der gesellschaftliche Nutzen als „selbstverständliche Folge“ von wissenschaftlichem Erkenntnisfortschritt angenommen (Schavan 2008, 9). Andererseits wird darauf hingewiesen, dass Leistungen die für die Erbringung eines gesellschaftlichen Nutzens notwendig sind (z. B. Dialog mit Praxis und Gesellschaft) in Zeit- und Ressourcenkonkurrenz zu Leistungen stehen, die für die etablierte Evaluierung relevant sind (z. B. Publikationen) (vgl. Stoll-Kleemann 2007).
- Für die Evaluierung inter- und transdisziplinärer Forschung gibt es zum einen vielfältige Einzelkonzepte, die auch auf den gesellschaftlichen Nutzen Bezug nehmen, jedoch bisher nicht breit angewendet werden. Zum anderen werden in verschiedenen Programmen der öffentlich geförderten Forschung Zielsetzungen wie gesellschaftlicher Nutzen, Wissenstransfer, und Beiträge zur nachhaltigen Entwicklung gefordert. Solche Leistungen der Forschung für Praxis und Gesellschaft müssen in Anträgen und Berichten der öffentlichen Forschungsförderung auch durchaus dokumentiert werden, (z. B. in Formulierung und Fortschreibung eines Verwertungsplanes bei Förderung durch Projektträger des BMBF, BMWi, BMFSFJ, BMELV und BMU (siehe Kapitel 5.1). Somit werden gewisse Anreize für die Erbringung von Leistungen für Praxis und Gesellschaft durch einzelne Institutionen sowie durch die Anforderungen einzelner Förderprogramme gesetzt. Starke Lenkungseffekte auf die Forschung wie durch die etablierten Performanzindikatoren (z.B. SCI) werden dadurch aber nicht erreicht.



Abbildung 1: Gesellschaftlicher Nutzen in Leitbildern öffentlich geförderter Forschung, in der etablierten Forschungsevaluierung und die Anreize für die Ausrichtung der Forschung.

1.3 Zielsetzung

Das Projekt soll dazu beitragen, dass künftig neben dem Scientific Impact in der Forschungsevaluierung verstärkt die Leistungen der Forschung für Praxis und Gesellschaft berücksichtigt und honoriert werden. Dadurch sollen Anreize zur Lösung aktueller gesellschaftlicher Probleme verstärkt werden (vgl. Abbildung 2). Denn der gesellschaftliche Nutzen, der vielfach ein zentrales Ziel öffentlich finanzierter Forschung ist, trat bisher in der Regel gegenüber der Konzentration auf wissenschaftliche Erkenntnis und dem Erreichen möglichst hoher wissenschaftlicher Reputation in den Hintergrund.

Die konkrete Zielsetzung des Projektes umfasst daher vier Aspekte, die hier am Beispiel der (ökologischen) Agrarwissenschaften bearbeitet wurden.

1. Kritische Analyse bestehender Evaluierungsinstrumente als Basis für eine veränderte Bewertung von Forschungsleistungen
2. Geeignete Kriterien für die (ökologischen) Agrarwissenschaften identifizieren (Literatur und Interviews)
3. Vorhandene Evaluierungsinstrumente analysieren und ergänzen
4. Etablierungsstrategie für eine ergänzende Evaluierung im Agrarsektor entwickeln

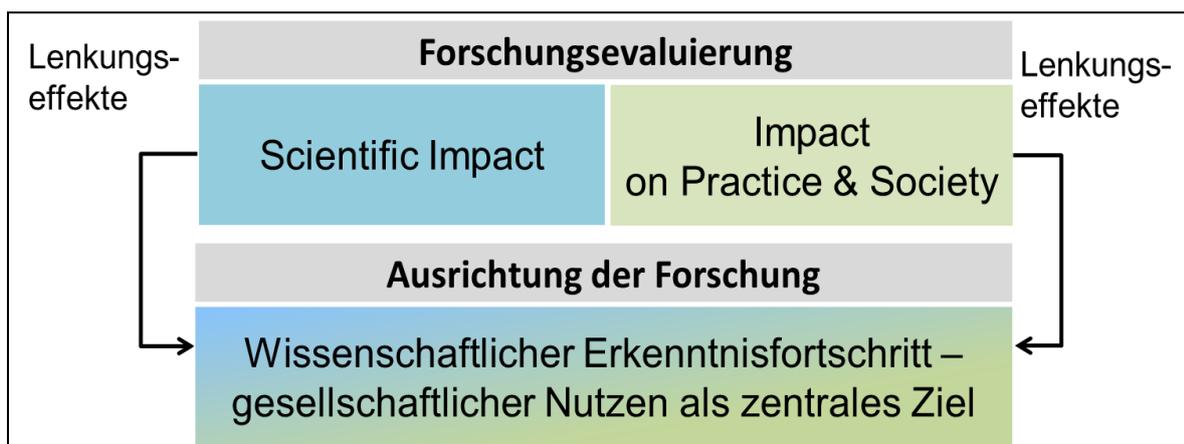


Abbildung 2: Langfristiges Ziel des Projektes: Forschungsevaluierung berücksichtigt den „Scientific Impact“ und den „Impact on Practice and Society“, mit entsprechenden Lenkungseffekten auf die Ausrichtung der Forschung.

Veränderung der Zielsetzung im Projektverlauf:

Dieses Forschungsvorhaben ging vom Ziel aus, basierend auf den Ergebnissen von Punkt 1-3, einen Leitfaden zur Bewertung angewandter, praxisorientierter und transdisziplinärer Forschung im Bereich der Ökologischen Landwirtschaft zu erarbeiten.

Aufgrund der Ergebnisse zu Punkt 1-3 wurde jedoch das 4. Teilziel, die Entwicklung einer Etablierungsstrategie für die Ergänzung der Forschungsevaluierung, ausgeweitet und in einer Projektverlängerung intensiv bearbeitet.

1.4 Begriffsklärung „Praxis“, „Gesellschaft“, „praxisorientiert“

Als Praxis wird die jeweilige (u. U. sehr weit gefasste) nicht-wissenschaftliche Zielgruppe der Forschung verstanden. Mit dem Begriff „Gesellschaft“ werden ergänzend auch andere Betroffene und deren Grundbedürfnisse bzw. das Gemeinwohl eingeschlossen. Da diese Begriffe sehr breit sind, müssen Zielgruppen und andere Betroffene für die Leistungen der Forschung unbedingt genauer spezifiziert werden. In der Landwirtschaft sind dies bspw. die Akteure entlang der gesamten Wertschöpfungskette von der Erzeugung bis zum Konsumenten, sowie Beratung, Verwaltung, Bildung, Politik, NGO's.

Der Begriff „praxisorientiert“ soll alle Forschungsformen einschließen, die ihre Leistungen an die Praxis oder die Gesellschaft adressieren, also zur Lösung bestimmter Probleme beitragen wollen – unabhängig davon, ob dies mit inter- oder transdisziplinärer Forschung oder einem klassischen Forschungsansatz erfolgt, bei dem erst nach Abschluss der Forschungsarbeiten im engeren Sinne ein Wissenstransfer in die Praxis erfolgt.

2 Projektablauf, Material und Methoden

2.1 Projektphase 1 - Entwicklung eines Leitfadens für die Evaluierung praxisbezogener Ökolandbauforschung

Für die Entwicklung von Vorschlägen zur Bewertung angewandter, praxisorientierter, inter- und transdisziplinärer Forschung wurden drei Kernthemen in drei methodischen Schritten im Zeitraum von 10/2008 – 03/2010 bearbeitet:

Kernthemen:

- a) Kritik am etablierten Evaluierungssystem, insbesondere aus Sicht angewandter, praxisorientierter, inter- und transdisziplinärer Forschung.
- b) Zusammenfassende Analyse bestehender Konzepte für die Evaluierung inter- und transdisziplinärer Forschung, jeweils mit dem Fokus auf die Nachhaltigkeits- und Agrarforschung.
- c) Erarbeitung von Vorschlägen zur Evaluierung praxisorientierter Ökolandbauforschung.

Methodische Schritte:

- 1) Literaturanalyse zu a) und b).
- 2) Leitfadengestützte qualitative Interviews mit 22 Experten der praxisorientierten Ökolandbauforschung zu a) und c).
- 3) Synthese von Interviewergebnissen und Literatur und Weiterentwicklung der Projektergebnisse, u. a. über einen Workshop (WS) mit Gruppendiskussion (sechs Interviewpartner und vier weitere Experten).

Die Interviews wurden mit einem explorativen Ansatz geführt, um eine Perspektivenverengung auf vorhandene Konzepte aus der Literatur zu vermeiden und zunächst die Ziele, Anforderungen und Vorschläge zur Evaluierung direkt im Feld der Ökolandbauforschung zu erheben (Lamnek 1995, 127). Dadurch konnten die Spezifika des Ökolandbaus angemessen bei der Analyse von Ergebnissen aus anderen Forschungsbereichen berücksichtigt werden.

Die Interviewpartner wurden gezielt anhand der folgenden Kriterien ausgewählt:

- Ausgewiesene Experten der (Ökologischen) Agrarwissenschaften und angrenzenden Disziplinen in entsprechender thematischer Breite (Boden, Pflanze, Tier, Technik, Ökonomie, Sozialwissenschaften).
- Erfahrung mit inter- und transdisziplinärer Forschung, Wissenstransfer, Forschungsmanagement und Forschungsevaluierung.
- Verschiedene Forschungsrichtungen: Universitäre Forschung, Ressortforschungseinrichtungen des Bundes und Forschung an privaten bzw. (halb-)öffentlichen Forschungseinrichtungen.

Der Interviewleitfaden umfasst offene Fragen zu den folgenden Inhalten:

- Beurteilung der etablierten Forschungsevaluierung
- angemessene Evaluierung von Ökolandbau- und Nachhaltigkeitsforschung, ausgerichtet an den Kriterienbereichen
 - o Praxisimpact,
 - o Interdisziplinarität,

- Systembezug,
 - Langfristig ausgerichtete Forschung,
 - Vielfalt von Themen und Methoden in der Forschung,
 - Gesellschaftliche Brisanz/Relevanz,
 - Originalität,
 - normative Kriterien, abgeleitet aus der Literatur zu den Prinzipien des Ökolandbaus (Lampkin 1990, Lindenthal, Vogl und Heß 1996, IFOAM 2008) als auch aus der Kritik am etablierten Evaluierungssystem (siehe Kap. 3.3).
- relevante Zielgruppen der Forschung und Evaluierung
 - freie Ergänzungsmöglichkeiten zu den gestellten Fragen

Abgeschlossen wurden die Interviews durch eine semiquantitative Bewertung von in der Literatur als besonders geeignet identifizierten Kriterien und Instrumenten (7 Fragen mit einer Skala von 1 „stimme voll“ zu bis 6 „stimme überhaupt nicht zu“).

Die Interviews wurden volltranskribiert, mit atlas.ti codiert und nach Methoden der qualitativen Sozialforschung, vorwiegend deskriptiv ausgewertet. (u.a. Atteslander 2000, 156–227, Flick 2002, 279).

Im nachfolgenden Text wird jeweils die Anzahl der Personen, die explizit zu einem Sachverhalt Vorschläge, Zustimmung oder Ablehnung geäußert haben, in Relation zur Grundgesamtheit der befragten Interviewpartner¹ (z. B. 6/22). bzw. zur Anzahl Workshopteilnehmer (WS) (z. B. 4/10 WS) dargestellt. Ergebnisse der semiquantitativen Bewertung werden nach demselben Prinzip dargestellt und sind mit dem Kürzel (Q) gekennzeichnet.

2.2 Projektphase 2 - Transfer und Diskussion von Projektergebnissen

Der Projektteil 2 umfasste eine Transferphase von 3 Monaten ab 04/2010, die kostenneutral bis 12/2010 verlängert wurde, um das Evaluierungskonzept weiter zu konkretisieren. Die Transferphase diente dazu Feedback des wissenschaftlichen und wissenschaftspolitischen Umfeldes in die erarbeiteten Projektergebnisse einzubauen. Diskutiert wurden Anwendbarkeit und Verbesserungsbedarf der Kriterien und Instrumente sowie Hindernisse, Potentiale und ggf. Umsetzungsschritte einer Integration der Vorschläge in die bestehende Evaluierungspraxis. Es wurden sechs Vortragsveranstaltungen mit Diskussion durchgeführt (siehe Kap. 10.2). Die Ergebnisse dieser Veranstaltungen wurden protokolliert und in die Weiterentwicklung des Konzeptes integriert.

2.3 Projektphase 3 - Konzept für die strukturierte Dokumentation von Leistungen der Forschung für Praxis und Gesellschaft in Anträgen und Berichten

Im Projektteil 1 und 2 wurden als entscheidendes Hemmnis für die Bewertung von Forschungsleistungen für Praxis- und Gesellschaft der Aufwand für deren Dokumentation identifiziert: In den vorhandenen Konzepten zur Evaluierung inter- und transdisziplinärer Forschung werden diese Leistun-

¹Die Grundgesamtheit ist nicht immer gleichbleibend, da einzelne Fragen von einigen wenigen Interviewpartnern u. a. aus Zeitgründen nicht beantwortet wurden.

gen häufig über Interviews und Dokumentenanalysen erfasst, womit ein erheblicher Aufwand für die Evaluation verbunden ist. Darüber hinaus werden diese Leistungen in der Regel jeweils nur für eine spezifische Evaluierung erfasst, mit der Folge, dass einmal erhobene Daten für andere Evaluierungen nicht verfügbar oder brauchbar sind. Im Projektteil 2 wurden daher Möglichkeiten gesucht, den Datenerfassungsaufwand zu verringern und die Dokumentation der Leistungen dauerhaft für verschiedene Evaluierungen verfügbar und aktuell zu halten.

- Deshalb wurden in einem ersten Schritt für die zur Evaluierung benötigten Kriterien weitgehend standardisierte Fragen und Antworten ausgearbeitet.
- Dieser Fragenkatalog bzw. Datenbedarf wurde mit den bestehenden Dokumentationsanforderungen für Anträge und Berichte der öffentlichen Forschungsförderung und den inhaltlichen Zielsetzungen einzelner Programme verglichen.
- Weitreichende Überschneidungen des Datenbedarfs der Forschungsförderung und der Evaluierung von Leistungen für Praxis und Gesellschaft legten die Erarbeitung eines Vorschlags nahe, wie die Datenerfassung für die Evaluierung schon bei der Beantragung und Berichterstattung effizient eingebunden werden kann.
- Mit VertreterInnen mehrerer Forschungsförderungseinrichtungen und ForscherInnen wurde in Workshops über die Möglichkeiten diskutiert, die Datenerfassung für die Evaluierung von Leistungen für Praxis und Gesellschaft in bestehende Dokumentationsprozesse für die Forschungsförderung zu integrieren. Ergänzend wurde erörtert, wie ein Gesamtkonzeptes für eine veränderte Forschungsevaluierung mittelfristig etabliert werden könnte.

3 Forschungsevaluierung: Einführung und Kritik

3.1 Einführung in die Evaluierung

3.1.1 Begriffe und Grundlagen der Evaluierung

Evaluation ist alltagssprachlich mit „Bewertung“ zu übersetzen und ist somit recht unspezifisch: *„Irgend etwas wird von irgend jemandem nach irgendwelchen Kriterien in irgendeiner Weise bewertet.“* (Kromrey 2001)

Eingrenzung und Präzisierung

Im wissenschaftlichen Sinne ist Evaluation dagegen die Anwendung von empirischen Methoden zur Informationsgewinnung und von systematischen Verfahren zur Informationsbewertung anhand offener Kriterien, die eine intersubjektive Nachprüfbarkeit möglich machen. Die Evaluierung soll dabei stets einen definierten Nutzen haben (Stockmann 2004). Dabei hat nach Kromrey (2001) der Nutzen der Evaluation im Konfliktfall einen höheren Stellenwert als die Wissenschaftlichkeit des Verfahrens. Auch ist Evaluation selbst keine Wissenschaft, sondern die Ergebnisse und Empfehlungen einer Evaluation sind (bestenfalls) durch wissenschaftlich fundierte Verfahren erzielte Befunde (Hornborstel 2010, 296). Neben der Abgrenzung zur Wissenschaft ist Evaluation aufgrund unterschiedlicher Funktionen (s.u.) gleichzeitig auch nur schwer von Managementtechniken inkl. Qualitätsmanagement, Controlling und Auditing abzugrenzen (Hornborstel 2010, 295). Eine professionelle Evaluation muss somit genauer spezifiziert werden.

Der Evaluierungsgegenstand

Nicht „Irgendetwas“ sondern ein bestimmter Evaluierungsgegenstand wird evaluiert. Im vorliegenden Projekt wird bei der Evaluation der Forschung zum Ökologischen Landbau zwischen Wissenschaftlern, Projekten, Institutionen und Forschungsprogrammen unterschieden. Aufbauend auf einer solchen Klassifizierung von Evaluierungsgegenständen kann noch eine weitere Präzisierung erfolgen, da jeder Evaluierungsgegenstand durch drei interdependente Dimensionen nämlich Ziele, Maßnahmen und Effekte gekennzeichnet ist, die jeweils von Umgebungseinflüssen betroffen sind (Kromrey 2001). Des Weiteren unterscheiden sich Evaluierungsgegenstände in Umfang und Komplexität (DeGEval 2008).

Wer evaluiert?

Nicht „irgendjemand“ führt eine Evaluation durch, sondern dafür als kompetent erachtete Personen. Dafür sind jedoch verschiedene Funktionen im Evaluierungsprozess analytisch klar voneinander zu trennen – Informationsbeschaffung, Evaluierung, Ableitung von Konsequenzen aus den Befunden – wofür meist auch verschiedene Personen bzw. Kompetenzen notwendig sind (Kromrey 2001, 110).

Eine grobe Unterscheidung besteht in der internen und der externen Evaluation. Die interne Evaluation (Programmdurchführende und Evaluatoren entstammen derselben Organisation) hat die Vorteile eines geringeren Aufwandes, hoher kontextbezogener Sachkenntnis der Personen und einer unmittelbaren Nutzbarkeit der Ergebnisse. Nachteile bestehen in einer ggf. mangelnden Methodenkompetenz und fehlender Distanz in der Beurteilung. Wird die Evaluation sogar direkt von den Evaluierten vorgenommen, handelt es sich um eine Selbstevaluation.

Externe Evaluatoren weisen in der Regel eine größere Unabhängigkeit und Methodenkompetenz sowie professionelles Evaluations- und Fachwissen auf. Auf externen Evaluationsergebnissen basie-

rende Reformvorschläge können größere Legitimität und Einflussstärke entfalten. Externe Evaluationen und ihre Ergebnisse können jedoch auch Angst und Abwehr bei den Evaluierten und Umsetzungsprobleme hervorrufen. (Stockmann 2007).

Zum Ausgleich der jeweiligen Stärken und Schwächen können interne und externe Evaluationen sinnvollerweise miteinander verschränkt werden (Stockmann 2007). Dieses Verschränken findet auch bei der Evaluierung von Forschungsleistungen statt (Hornborstel 2010, 302) und wird insbesondere für problemorientierte und inter- und transdisziplinäre Forschung empfohlen (Schiller et al. 2005, Blackstock, Kelly und Horsey 2007, Walter et al. 2007, Defila, Di Giulio und Scheuermann 2008).

Mit welchen Kriterien wird evaluiert?

Es wird nicht mit „irgendwelchen Kriterien“ evaluiert, sondern es besteht der Anspruch, mit für den Evaluierungsgegenstand spezifischen und auf die Ziele der Evaluierung abgestimmten Kriterien zu arbeiten. Dabei stellt sich die Frage durch wen und in welcher Weise Kriterien definiert werden. Erste Unterscheidungen lassen sich hinsichtlich qualitativer und quantitativer Kriterien treffen (Kromrey 2001), aber auch dahingehend welche Dimension des Evaluationsgegenstandes vorrangig betrachtet wird (Kontext, Struktur, Konzept, Input, Prozess oder Wirkung) (DeGEval 2008).

Kriterien sollen dabei verständlich (für Wissenschaft, Praxis und Evaluatoren), einfach erfassbar bzw. messbar und leicht zu interpretieren sein. Dabei sollen sie relevant für die Entscheidungsträger sein und eine hohe Akzeptanz besitzen, wofür es erforderlich ist, dass sie eng verbunden mit den Zielen des Evaluierungsgegenstandes sind. Nicht zuletzt ist es erforderlich, dass die Kosten für die Datengewinnung in einem angemessenem Verhältnis gegenüber dem Nutzen der Information stehen (Schmid O. et al. 2008, 78f).

Wie wird evaluiert?

Evaluation „in irgend einer Weise“ ist in der professionellen Evaluierung ein objektiviertes Verfahren unter Verwendung von (überwiegend) Methoden der empirischen Sozialforschung zur Informationsbeschaffung und Bewertung. Dabei kann mit quantitativen oder qualitativen Methoden gearbeitet werden (Kromrey 2001). Wie evaluiert wird hängt vom Ziel bzw. Zweck der Evaluierung ab (s.u.). Je nach Evaluationskonzept und Zeitpunkt kann eine Evaluation formativ (aktiv-gestaltend, prozessorientiert, konstruktiv und kommunikationsfördernd) oder summativ (zusammenfassend, bilanzierend und ergebnisorientiert) ausgerichtet sein. Geeignet ist der formative Ansatz für eine Evaluierung ex ante (im Vorfeld) oder on-going (begleitend) und der summative Ansatz für die on-going oder ex post Evaluierung (nach Abschluss) (Stockmann 2004).

Welches Ziel bzw. welchen Zweck hat die Evaluierung?

Chelimsky (1997, 100 ff. in Kromrey 2001, 113-15) unterscheidet drei „conceptual frameworks“ der Evaluation: (a) zur Verbreiterung der Wissensbasis (Forschung), (b) zu Kontrollzwecken und (c) zu Entwicklungszwecken. Zum Zweck der Forschung ist die Evaluierung vornehmlich eine wertneutrale Wirkungsanalyse, die Kontrolle impliziert eine quantitative, summative, ex post Betrachtung, wohingegen der Zweck der Entwicklung besonders gut durch eine formative Evaluierung erreicht wird (Kromrey 2001, 113-15).

Stockmann (2004) unterscheidet dagegen zwischen vier untereinander interdependenten Zielen/Zwecken: (a) die Gewinnung von Erkenntnissen, (b) die Ausübung von Kontrolle, (c) die Schaffung

von Transparenz, um einen Dialog zu ermöglichen (Lernfunktion) und (d) die Dokumentation des Erfolgs (Legitimation).

Welche Aufgaben kann eine Evaluierung haben?

Das Aufgabenprofil einer Evaluierung wird von Stockmann (2004) in fünf Hauptrichtungen differenziert:

- a) Evaluierung von Ablaufprozessen: z. B. Potenziale und Probleme bei der Implementierung von Programmen². Beurteilung, ob das Programm zielführend ist, anhand einer Beurteilung der Einhaltung des Zeitrahmens, Akzeptanz der Maßnahmen bei den Betroffenen, Interessenskonflikte, Ressourcen (finanziell und personell), Kommunikation, Koordination, Maßnahmen.
- b) Überprüfung der Zielerreichung „Soll-Ist-Vergleich“: Diese Aufgabe ist besonders schwer zu erfüllen, da Ziele oft nur sehr verschwommen formuliert werden, dokumentierte Ziele und tatsächliche Ziele voneinander abweichen, Ziele im Zeitverlauf veränderbar sind und verschiedener Akteure auch unterschiedliche Ziele haben.
- c) Evaluierung von Wirkungen: Diese sind zwischen intendierten und nicht-intendierten und positiven und negativen Wirkungen zu differenzieren und bedürfen einer Kausalitätsprüfung.
- d) Relevanz-/Signifikanzprüfung stellen die Ziele eines Programms selbst in Frage und prüfen, ob die Ziele und Maßnahmen überhaupt geeignet sind, um relevante Entwicklungs- oder Innovationsleistungen zu erbringen.
- e) Mittels Kausalitätsprüfung wird festgestellt, ob die ermittelten Wirkungen überhaupt dem Programm oder aber externen Einflüssen zuzuschreiben sind. Dieses wird als eine der schwierigsten Evaluationsaufgaben bezeichnet. (Stockmann 2004).

3.1.2 Standards für die Evaluation

Um die Qualität von Evaluationen zu fördern wurden allgemeine Standards der Evaluation definiert, die aus den „Standards for Evaluation“ des Joint Committee on Standards for Educational Evaluation abgeleitet sind (sie gelten nur bedingt für Selbstevaluationen) (Stockmann 2004).

Diese Standards der Evaluation umfassen 25 Einzelkriterien, zu vier geforderten grundlegenden Eigenschaften von Evaluation: Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit.

- a) Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszwecken sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausrichtet.
- b) Die Durchführbarkeitsstandards sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt wird.
- c) Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen wird.
- d) Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt. (DeGEval 2008, 10-13).

² Programm meint hier allgemein ein Vorhaben, welches durch Ziele und Zeitpunkte spezifiziert werden kann.

3.1.3 Erkenntnistheoretische Fundamente der Evaluation

Die aufgeführten pragmatischen Unterschiede in der Evaluierung sind zu großen Teilen auf verschiedene erkenntnistheoretische Ansätze der Evaluation zurückzuführen, die hier nach Stockmann (2007) kurz dargestellt werden.

Positivistischer Ansatz

Der positivistische Ansatz geht davon aus, dass es eine einzige Realität gibt, die objektiv erfassbar ist. Evaluation ist ein empirisch-wissenschaftliches Verfahren mit Erkenntnis und Verwertungsinteressen mit besonderem Nutzen der Ergebnisse für die Praxis. Objektivität und Neutralität sind entscheidende Standards, der Schwerpunkt liegt darauf, Ursache-Wirkungszusammenhänge zu messen und die Generalisierung von Ergebnissen zu ermöglichen.

Konstruktivistischer Ansatz

Der konstruktivistische Ansatz geht davon aus, dass es keine „wahre“ Realität gibt, die mit empirisch-wissenschaftlichen Verfahren erfasst werden kann. Realität ist aus verschiedenen Perspektiven sozial konstruiert, die in Konflikt zueinander stehen können. Mit qualitativen Methoden wird versucht, den Evaluierungsgegenstand im jeweiligen Kontext zu erklären und zu verstehen.

Transformativer, emanzipativer Ansatz

Aufbauend auf dem konstruktivistischem Ansatz wird berücksichtigt, dass verschiedene Realitäten einen unterschiedlichen Einfluss auf den Kontext des Evaluierungsgegenstandes haben. Methoden und Ergebnisse sollen diese verschiedenen Perspektiven nebeneinander verdeutlichen. Evaluatoren sollen mit den Beteiligten und Betroffenen interagieren, (ohne die Standards der Evaluation zu beeinträchtigen), mit dem Ziel die Verhältnisse zugunsten der Benachteiligten zu verändern.

Eine Annäherung dieser verschiedenen Paradigmen zeigt sich darin, dass unter Evaluierungsexperten ein weitgehender Konsens dahingehend besteht, dass Evaluation die Perspektiven und Bedürfnisse der Stakeholder berücksichtigt, qualitative und quantitative Methoden miteinander kombiniert werden (Multimethodenansätze) und einen Nutzen für Auftraggeber und Stakeholder generiert werden soll.(Stockmann 2007).

3.1.4 Diskussion des eigenen Evaluierungsparadigmas

Im vorliegenden Projekt wird vornehmlich – jedoch nicht ausschließlich – auf eine summative ex-post Evaluierung fokussiert, um die Zielsetzung des Praxis-Impact Projektes, nämlich eine angemessene Honorierung von erbrachten Leistungen der Wissenschaft für Praxis und Gesellschaft, zu erreichen. Zentrale Aufgabe des vorliegenden Projektes ist daher, ein vielfältiges Kriterienset für eine kontextbezogene Auswahl für verschiedene Evaluierungsgegenstände zu erarbeiten.

Dabei ist es von außerordentlicher Bedeutung, dass durch das zu definierende Kriterienset positive Lenkungseffekte (hier für die Weiterentwicklung von Forschung und Praxis des Ökologischen Landbaus) erreicht werden können.

Die Frage welche Informationen auf welche Weise erfasst werden können, um inhaltlich als notwendig erachtete Kriterien für verschiedene Evaluierungsgegenstände adäquat bewerten zu können, ist eine wichtige Aufgabe des vorliegenden Projektes.

Dabei stehen die Projektbearbeiter dem konstruktivistischen Paradigma näher, als dem positivistischen. Dazu gehört, dass auch im Hinblick auf eine Evaluation das „Nichtwissen“ über die Zukunft akzeptiert wird und ein „Irrtum der Evaluation“ stets möglich ist bzw. dass grundsätzlich immer nur eine Annäherung an die Realität gelingen kann.

Der Zweck einer Evaluierung sollte daher den Schwerpunkt haben, Lerneffekte zu generieren (vgl. Böcher und Krott 2010), die intrinsische Motivation zu stärken und einen Beitrag zu ‚lernenden Organisationen‘ zu leisten (vgl. Frey 2010 in Böcher und Krott 2010).

Konsequenzen aus einer Evaluierung sollten sich demzufolge auch eher an einer Verbesserung der Voraussetzungen für die Zielerreichung orientieren.

3.2 Einführung in die Forschungsevaluierung

Die Evaluierung von Forschungsleistungen basiert zurzeit im Wesentlichen auf der Publikation von Ergebnissen in Zeitschriften mit Peer-Review und der Akquise von Drittmitteln. Ergänzend werden anlassbezogene Evaluierungen (z. B. von Instituten) durchgeführt, wofür u. a. die Ergebnisse und Indikatoren aus den beiden erstgenannten Elementen der Forschungsevaluierung herangezogen werden.

3.2.1.1 Publikationen in Journals mit Peer-Review

Der derzeit wichtigste Teil der Forschungsevaluierung wird in Verbindung mit der Veröffentlichung von Ergebnissen vollzogen und ist innerhalb des Wissenschaftsystems hochgradig zentralisiert und institutionalisiert.

Relevant für die Forschungsevaluierung sind wissenschaftliche Zeitschriften mit einem Peer-Review Prozess als qualitative Komponente des Evaluierungssystems. Voraussetzung für diese Art der Evaluierung ist die Bereitschaft von Forschern, sich als Gutachter im Rahmen des Peer-Review (ehrenamtlich und kostenfrei) zu engagieren und für das Bewerten der Artikel von Kollegen entsprechend Zeit aufzuwenden.

Diese Zeitschriften sind überwiegend englischsprachig und werden im System von Thompson Reuters, ehemals ISI³, geführt. Über Thompson Reuters wird in einem jährlichen Journal Citation Report zentral erfasst, wie häufig eine Publikation von anderen Wissenschaftlern in Zeitschriften mit Peer Review zitiert wird, wodurch weitere wichtige Indikatoren, der etablierten Evaluierung generiert werden, nämlich die Citation Indices (für Publikationen, Autoren und Institutionen) und der Journal Impact Faktor um die Journals zu gewichten. Details zu diesen Indikatoren sind in Tabelle 1 aufgeführt.

³ ISI: Institute of Scientific Information in Philadelphia / USA

Tabelle 1: Glossar zu den wichtigsten zitationsbasierten Indikatoren der etablierten Forschungsvaluierung

Citation Indices des JCR (allg.)	Analysiert wird grundsätzlich wie viele andere Artikel eine Publikation zitieren. Citation Indices, wie auch H-Index und JIF basieren auf dem Journal Citation Report und beziehen grundsätzlich nicht nur Zitationen in Artikeln, sondern auch „letters, corrections and retractions, editorials, and other items“ mit ein. Sie sind für verschiedene Bereiche verfügbar und über Thompson Reuter grundsätzlich kostenpflichtig. http://thomsonreuters.com/products_services/science/free/essays/concept_of_citation_indexing/
SCI	Science Citation Index Expanded. Wissenschaftliche Zitationsdatenbank von Thompson Reuters. Schließt >8300 Zeitschriften mit Peer Review über 150 Disziplinen ein. (http://thomsonreuters.com/products_services/science/science_products/a-z/)
SSCI	Zitationsdatenbank von Thompson Reuters für die Sozialwissenschaften. Schließt >4.500 Fachzeitschriften mit Peer Review über 50 Disziplinen ein.(http://thomsonreuters.com/products_services/science/science_products/a-z/)
A&HCI	Arts and Humanities Citation Index >2.300 Fachzeitschriften (und ausgewählte Teile von 6000 SCI und SSCI-Zeitschriften) (http://thomsonreuters.com/products_services/science/science_products/a-z/)
H-Index	H-Index = N Datensätze mit mindestens N Zitaten reduziert die überproportionale Gewichtung von hoch bzw. gar nicht zitierten Publikationen, Zeitspanne der Auswertung kann individuell gewählt werden, entwickelt durch J.E. Hirsch (http://images.webofknowledge.com/WOK45/help/WOS/h_citationrpt.html)
JIF	Journal Impact Faktor basiert auf dem jährlichen Journal Citation Report (JCR) (>10.000 Journals) und beschreibt die Häufigkeit mit der die Artikel eines Journals im Durchschnitt in anderen Journals in einem Jahr (oder einer definierten Periode) zitiert werden. Der jährliche JCR Impact Factor eines Journals, z.B. für 2010 wird folgendermaßen berechnet: $\frac{\text{Zahl der Zitate in 2010 aus den Artikeln von 2008 und 2009 eines Journals}}{\text{Zahl der Artikel von 2008 und 2009 eines Journals}}$ Es gibt aber auch einen fünf-Jahres Impact Faktor und Selbstzitationen (innerhalb des Journals) können eliminiert werden. http://thomsonreuters.com/products_services/science/free/essays/impact_factor/

3.2.1.2 Drittmittelakquisition

Die Drittmittelvergabe und damit die Kriterien und der Begutachtungsprozess von Forschungsanträgen sind dagegen je nach Förderinstitution unterschiedlich gestaltet. Die Drittmittelakquisition ist für Forscher in der Regel mit hohem Zeitaufwand für die Erstellung von Skizzen und Anträgen verbunden. Nach Projektbeauftragung sind die Forschungsleistungen in der Regel in Berichten ausführlich zu dokumentieren. Für die Evaluierung von Wissenschaftlern und Institutionen wird aus der Drittmittelakquisition derzeit im Wesentlichen nur ein Kriterium genutzt: das Drittmittelvolumen. Dieses wird häufig differenziert nach einer unterschiedlichen „Wertigkeit“ der Drittmittelgeber sowohl in Hinblick auf ihr Gutachterverfahren, als auch ihrer Zielsetzung. Sogenannte „qualifizierte Drittmittel“ z.B. vergeben durch die DFG werden dabei höher bewertet, als die Forschungsförderung durch Ministerien.

3.2.1.3 Anlassbezogene Evaluierung

Die dritte entscheidende Dimension der Forschungsvaluierung ist die anlassbezogene Evaluierung. Individuelle Forschungs- und Lehrleistungen von Wissenschaftlern werden für organisationsinterne Evaluierungen (Berufungen, leistungsorientierte Mittelvergabe) aber auch aggregiert für die Evaluierung von Institutionen herangezogen. Als handlungsfähige Organisationseinheiten haben sie auch die Fähigkeit, Evaluierungsempfehlungen umzusetzen. Evaluierungen von Förderprogrammen haben die Erreichung von Förder-/Programmzielen zum Inhalt. Aber auch Arbeitsgruppen werden evaluiert. Querschnittsevaluierungen befassen sich mit fachübergreifenden Forschungsschwerpunkten oder

disziplinären Forschungsleistungen und -bedingungen. System- und Strukturevaluierungen gehen noch darüber hinaus und evaluieren die institutionellen Strukturen des Wissenschaftssystems bzw. ihrer Funktionsbereiche (Begutachtungswesen, Qualitätssicherung, Steuerung). (Hornborstel 2010, 302).

Gerade in der Wissenschaftsevaluierung sind fließende Übergänge oder Verknüpfungen zwischen interner und externer Evaluierung festzustellen, indem z. B. interne Audits für eine externe Evaluierung verwendet werden oder die universitäre Selbstverwaltung Einfluss auf die Auswahl von Evaluatoren, Kriterien und die Verwendung der Ergebnisse haben (Hornborstel 2010, 302).

Die beschriebenen anlassbezogenen Evaluierungen sind oft mehr oder weniger spezifisch auf den Evaluierungsgegenstand und seine Zielsetzungen bezogen. Gleichzeitig sind die Indikatoren der Publikationsleistung und das Drittmittelvolumen nach wie vor die dominanten Kriterien für die Beurteilung wissenschaftlicher Leistungen, oft ergänzt durch weitere quantitative Indikatoren, wie die Anzahl an Doktorarbeiten und Habilitationen. Darüber hinaus haben die Ergebnisse einer anlassbezogenen Evaluierung häufig Einfluss auf die Finanzierung von Institutionen, z. B. im Rahmen der leistungsorientierten Mittelvergabe (LOM) an Universitäten in Deutschland (Herb 2008).

In der Wissenschaft gibt es somit einerseits etablierte Verfahren, (Peer-Review von Publikationen, Begutachtung von Forschungsanträgen für die Projektförderung, Berufungsverfahren für Professoren). Andererseits ist in der Wissenschaft die Evaluation besonders umstritten, methodisch herausfordernd und in ihren Folgewirkungen unklar (Hornborstel 2010, 296f), worauf im Folgenden übersichtsartig eingegangen werden soll.

3.3 Das etablierte Evaluierungssystem in der kritischen Diskussion

3.3.1 Kritikpunkte am etabliertem Evaluierungssystem in der Literatur

Die etablierte wissenschaftliche Evaluierung basiert auf dem Peer-Review und quantitativen Kriterien, wie der Anzahl an Publikationen (mit Peer Review), Zitationszahlen, Journal-Ranking und Drittmittelvolumen. Dieses innerhalb der Wissenschaft etablierte Vorgehen bei der Evaluierung setzt Lenkungseffekte und Anreizstrukturen, die zunehmend kritisch bewertet werden. Die Diskussion über Voraussetzungen und Folgen outputorientierter Anreizsysteme für die universitäre und außeruniversitäre Forschung, die auf den etablierten Performanzindikatoren basieren, wird bereits in wissenschaftlichen Gemeinschaften (z. B. dem Wissenschaftsrat) geführt. Sie sind bspw. Inhalt einer aktuellen Arbeitsgruppe des Wissenschaftsrates, zu der derzeit jedoch noch keine Ergebnisse vorliegen (Wissenschaftsrat 2010, 10).

Ergänzungs- und Verbesserungsvorschläge sowie Kritik an den etablierten Kriterien und Verfahren der Forschungsevaluierung in der wissenschaftlichen Diskussion können folgendermaßen zusammengefasst werden (eine ausführliche Aufarbeitung ist im Anhang I dokumentiert):

- 1) Am etablierten Evaluierungssystem wird u. a. kritisiert, dass Wissenschaft sich dabei selbst bewertet (Fröhlich 2003a, Hornbostel 1999). Mit dem Focus auf der wissenschaftlichen Qualität bleiben dabei Fragen der gesellschaftlichen Relevanz, bzw. des gesellschaftlichen Nutzens der Forschung vielfach unbeantwortet. Als Folge wird befürchtet, dass die Verantwortung der Wissenschaft und Forschung gegenüber der Gesellschaft (u. a. im Hinblick auf die nachhaltige Ent-

wicklung) immer weniger als Bezugsrahmen dient. Stattdessen bestimmt das etablierte Evaluierungssystem - und damit die wissenschaftliche Community - verstärkt die Forschungspolitik (Stock 1994, Frank-Rieser 1999, Schlinghoff und Backes-Gellner, 2002, Fröhlich 1999, 2003a, 2003b, Brown 2007).

- 2) Quantitativ ausgerichtete Indikatoren der Forschungsevaluierung (v. a. Anzahl an Peer-Review Publikationen, Zitationszahlen, Journal-Ranking, Drittmittelvolumen) werden zunehmend kritisch bewertet, weil die inhaltliche Qualität und die gesellschaftliche Relevanz dadurch in den Hintergrund tritt. Sie führen statt dessen nach Ansicht einiger Autoren zu einem Zwang zum Publizieren bis hin zu einer „wissenschaftlichen Planwirtschaft“ (auf Basis von „Produktivitätskennziffern“) und hat eine zunehmende Unüberschaubarkeit des Wissens, Redundanzen sowie zunehmende Spezialisierung zur Folge (Frank-Rieser 1999, Fröhlich 1999, 2003b, Beisiegel et al. 2011).
- 3) Am etablierten Evaluierungssystem üben zahlreiche Autoren Kritik an dessen Logik und Umsetzung. Forschungsevaluierung ist durch das Peer Review und die innerhalb der wissenschaftlichen Community generierten quantitativen Indikatoren (z. B. SCI/SSCI/JIF), auch ein soziales System, mit spezifischen Wachstums- und Selektionsmechanismen sowie teilweise sich selbst verstärkenden Machtstrukturen. Als mögliche Folgen werden deshalb aufgeführt:
 - a) Die Vielfalt und Originalität von Themen und Methoden nimmt ab (Frey 2006, Nagel 2007), da große wissenschaftlichen Teil-Gesellschaften über Journals mit hohem JIF verfügen und damit auch die Relevanz und Akzeptanz von Themen und Methoden entsprechend beeinflussen (Schlinghoff und Backes-Gellner, 2002, Fröhlich 1999, 2003a und 2003b, Brown 2007). Darüber hinaus werden disziplinäre Paradigmen verfestigt und die „Normalwissenschaften“ gefördert (Fröhlich 2003a und 2003b) bzw. konservative Bewertungsmaßstäbe im Peer Review aufrechterhalten (Neidhardt 2010). Inter- und transdisziplinäre Forschungsansätze, aber auch ganz neue Fachgebiete und Fragestellungen müssen sich für Publikationen, wie auch für die Drittmittelakquisition, einem disziplinär ausgerichteten Peer Review stellen, dessen Qualitätsverständnis oft so eng ist, das neue Themen und Methoden nicht akzeptiert werden (s. Frank-Rieser 1999, Fröhlich 2003b, Ziegler 2003, Kueffer et al. 2007, Pohl und Stoll-Kleemann 2007, Albrecht et al. 2008). Dies betrifft auch wichtige Bereiche der Ökologischen Landwirtschaft und Nachhaltigkeitsforschung deren systemarerer Zugang häufig die Paradigmen der klassischen Wissenschaften sprengt (Lindenthal et al. 1996).
 - b) Die Förderung der Spezialisierung und Monodisziplinarität durch die disziplinäre Fragmentierung der Scientific Community und ihres Evaluierungssystems erfolgen dabei trotz immer komplexerer Fragestellung in ökologischen und gesellschaftlichen Bereichen, die interdisziplinäre Herangehensweisen erfordern würden (Frank-Rieser 1999, Fröhlich 2003b, Kueffer et al. 2007, Pohl und Stoll-Kleemann 2007, Albrecht et al. 2008)
 - c) Sich selbst verstärkende Prozesse sind in der Forschungsevaluierung systemimmanent. Z.B. entsteht eine Überbewertung von Journalen mit dem höchsten JIF (Young et al. 2008) oder sogenannter Hot-Papers u.a., da durch die begrenzte Anzahl an Publikationsorganen eine künstliche Knappheit entsteht. Die quantitativen Citation Indices begünstigen die Förderung von Zitierzirkeln sowie einseitige Ausrichtungen im Gutachtersystem und den jeweiligen wissenschaftlichen Gesellschaften und stärken die Macht großer wissenschaftlicher Communitys

(Stock 1994, Frank-Rieser 1999, Schlinghoff und Backes-Gellner, 2002, Fröhlich 1999 und 2003b, Brown 2007).

- d) Insgesamt ist ein zunehmender Konkurrenzdruck zu beobachten, der eine gezielte Anpassung an quantitative Parameter forciert (Frank-Rieser 1999, Nagel 2007). In vielen Bereichen der Forschungsförderung werden diese Entwicklungen noch befördert, indem jenen Forschungsinstitutionen und Arbeitsbereichen, die nach den etablierten Performanzindikatoren besonders erfolgreich sind, verstärkt Forschungsmittel zur Verfügung gestellt werden. Universitäten folgen dieser Entwicklung durch die leistungsorientierte Mittelvergabe (LOM) (Herb 2008). Besonders kritische Stimmen bringen das etablierte Evaluierungssystem daher auch mit einer Konzentration von Macht in Verbindung: Hohe Bekanntheit, viele Geräte und Forschungsgelder (und Macht) bringen somit noch mehr Bekanntheit, Geräte, Gelder und Beziehungen (Frank-Rieser 1999, Fröhlich 2003b). Eine Parallele zu vorherrschenden Paradigmen der Ökonomie und dem Wettbewerbssystem drängt sich zwangsläufig auf (Nagel 2007).
 - e) Irrtümer oder mangelhafte Sorgfalt von Gutachtern, haben negative Konsequenzen für den wissenschaftlichen Erkenntnisfortschritt und die Reputation von Wissenschaftlern oder Institutionen (vgl. z.B. Gans und Shepherd 1994).
 - f) Auch im etablierten Evaluierungssystem sind (wie in anderen Bewertungssystemen) Fälschungen und Plagiate möglich (Fröhlich 1999, 2003a und 2003b, Neidhardt 2010, Beisiegel et al. 2011). Wissenschaftliches Fehlverhalten wird durch zunehmenden Konkurrenzdruck und quantitativ ausgerichtete Indikatoren jedoch forciert (Nagel 2007, 3f.)
- 4) Die gängige Evaluierungspraxis in der Forschung hat trotz zunehmender Kritik in vielen Bereichen an Bedeutung gewonnen, was zur Folge hat, dass der „Scientific Impact“ die zentrale Qualitätsgröße darstellt. Dies benachteiligt jene Bereiche der angewandten Forschung, die sich auf einen hohen gesellschaftlichen Nutzen (v. a. in der Agrarforschung sowie in der Nachhaltigkeitsforschung) konzentrieren. Im Zuge der Exzellenzdiskussionen wird diese Art von Forschung sogar ernsthaft in Frage gestellt. Eine Benachteiligung besteht konkret in den folgenden Punkten:
- a) Eine unzureichende Berücksichtigung der, auf die jeweilige Landessprache orientierten, praxisbezogenen Forschung durch den fast ausschließlichen Einbezug englischsprachiger Zeitschriften (Fröhlich 2003a und 2003b, van Raan et al. 2011).
 - b) Abwertung inter- und transdisziplinärer, Paradigmen hinterfragender Forschung sowie neuer Themen und Methoden (siehe 3)a).
 - c) Leistungen außerhalb der wissenschaftlichen Publikation werden vom etablierten Evaluierungssystem nur unzureichend berücksichtigt, wie z.B. Publikationen für Zielgruppen in der Praxis oder Gesellschaft, Kooperationen oder Workshops mit Praxisakteuren, Beratungstätigkeiten oder die Lehre (Frey 2006, p.4f). Kommunikation innerhalb interdisziplinärer Projekte oder mit außerwissenschaftlichen Akteuren ist jedoch u. a. eine wichtige Voraussetzung für adäquate Lösungsbeiträge in der Nachhaltigkeitsforschung (Ziegler 2003), wie auch in der Forschung zum Ökologischen Landbau.
 - d) Interdisziplinäre Kooperation wird darüber hinaus unter einem steigenden Konkurrenzdruck zunehmend erschwert (Frank-Rieser 1999).

- 5) Gerade zu häufige Evaluationen führen zu Anreizverzerrungen in Richtung eines kurzfristigen Ergebnispotentials und verstärken risikoarme und outputorientierte Forschung. Schlimmstenfalls wird sogar die intrinsische Motivation reduziert. Darüber hinaus verursachen sie zunehmende Kosten und binden Zeitressourcen in der Forschung (Day und Klein 1990 zit. in Frey 2006, Imboden 2006, Frey 2006, Power 1997).
- 6) Einige Autoren befürchten, dass sich die Politik der etablierten Evaluierungswerkzeuge bedient, um sich bestimmten, wenig etablierten, wirtschaftspolitisch unbequemen Institutionen (die zudem meist per se auch geringeres privatwirtschaftliches Drittmittelvolumen aufweisen) zu entledigen (vgl. Ziegler 2003, Münch 2006 und 2007).
- 7) Nicht zuletzt wird kritisiert, dass die Forschungsevaluierung auf einem positivistischen, kausal-analytischen und stark quantitativ ausgerichteten Wissenschaftsverständnis basiert, welches nicht hinreichend reflektiert wird (Pietschmann 1980 a, b, Ziegler 2003).

Schlussfolgernd kann aus der aktuellen Debatte über die etablierten Evaluierungssysteme abgeleitet werden, dass die gängigen Evaluierungskriterien und -instrumente für angewandte Forschungsfelder zur Qualitätsbeurteilung unzureichend sind, insbesondere dann, wenn der „Impact für Praxis und Gesellschaft“ von Forschung bewertet werden soll. Diese fehlende Eignung betrifft in besonderem Maße „kleine“ Scientific Communities mit einer großen disziplinären Vielfalt und einem Anwendungsbezug der sich weniger in verwertbaren Produkten für Industrie, Handel und Gewerbe niederschlägt, sondern einen hohen Bezug zu öffentlichen Gütern aufweist, wie es in der Ökolandbauforschung der Fall ist.

Eine Ergänzung der etablierten Evaluierung um Kriterien, welche die Leistungen der Forschung für Praxis, Gesellschaft und Umwelt abbilden, erscheint daher dringend notwendig. Wie diese Leistungen dokumentiert und unter Einbindung von Anwendungspraxis und betroffenen Interessengruppen evaluiert werden können, ist deshalb zentrales Ziel dieses Forschungsvorhabens.

3.3.2 Einschätzung des etablierten Evaluierungssystems in den Interviewergebnissen

Im vorliegenden Projekt wurden die Interviewpartner um eine Einschätzung des etablierten Evaluierungssystems aus Sicht der Ökolandbauforschung gebeten. Die Interviewergebnisse (Tabelle 2) zeigen, anhand der Häufigkeiten der frei geäußerten Kritikpunkte, dass vielfach die gleichen oder sehr ähnlichen Kritikpunkte wie in der Literatur aufgeführt werden. Somit werden sie offensichtlich von den befragten Experten auch für die Forschung zum Ökologischen Landbau als relevant eingeschätzt. Von besonderer Wichtigkeit sind dabei die folgenden Kritikpunkte: fehlender Bezug zum Gesellschaftlichen Nutzen, Überbewertung quantitativer Indikatoren, Einschränkung der Vielfalt von Themen und Methoden, fehlende Berücksichtigung von Leistungen außerhalb der wissenschaftlichen Publikation (Vorträge, allgemeinverständliche Publikationen, etc.) sowie die Folgen von zu häufigen Evaluierungen.

Tabelle 2: Zusammenfassung der wichtigsten Kritik am etablierten Evaluierungssystem aus Sicht der Interviewpartner des Projektes.

Stichworte aus der Kritik in der Literatur, gegliedert entsprechend Kap. 3.3	Interviewergebnisse: Häufigkeit der Nennungen und wichtige ergänzende Aussagen
<p>1) „Innerwissenschaftlicher Fokus“ Gesellschaftlicher Nutzen und gesellschaftliche Verantwortung sind keine Bezugsgrößen</p>	<p>8/22 Wissenschaft bewertet sich selbst, Praxisimpact wird nicht erfasst</p>
<p>2) Quantitäten werden überbewertet, Qualität und Inhalt tritt in den Hintergrund</p> <p>Zwang zum Publizieren Redundanzen häufen sich Drittmittelvolumen ist zu quantitativ</p>	<p>11/22 „serielle Forschung“, Veröffentlichung nicht um Erkenntnis zu publizieren, sondern um Impact Faktor zu erhöhen, Etabliertheit, Bekanntheit und Namen für Bewertung, bzw. Veröffentlichung wichtiger als Qualität</p> <p>1/22 2/22</p> <p>8/22 viel Geld ausgeben ≠ viel „ehrenhaftes“ tun, Vernachlässigung einer Bewertung der „Relation“ zum Gesamtvolumen der für das Fach verfügbaren Drittmittel</p>
<p>3) Spezifische Selektions- und Wachstumsmechanismen</p> <p>3a) Vielfalt und Originalität von Themen und Methoden nimmt ab, Abwertung origineller, Paradigmen brechender Forschung</p> <p>3 b) Auf Spezialisierung und Monodisziplinarität ausgerichtet, Interdisziplinarität wird beeinträchtigt</p> <p>3 c) Sich selbst verstärkende Prozesse und Machtstrukturen Überbewertung von Zeitschriften mit hohem JIF und Hot-Papers geschlossene Gesellschaften, Zitationszirkel/Machtkartelle Wachsender Einfluss großer Gesellschaften, Benachteiligung von kleinen, neuen Fachbereichen</p> <p>3 d) Wachsender Konkurrenzdruck, Anpassung an quantitative Parameter</p>	<p>10/22 Einseitige (wissenschaftliche) Zusammensetzung der Peers, Einseitiger Fokus auf Grundlagenforschung/Methodenentwicklung</p> <p>4/22 spezialisierte Journals, wenige (niedrigererankte) Journals für inter- und transdisziplinäre Forschung und Ökologischen Landbau</p> <p>1/22 5/22 3/22</p> <p>4/22 „Anpassungen“ und „Pervertierung“</p>
<p>4) Benachteiligung praxisorientierter, inter und transdisziplinärer Forschung</p> <p>4 b) S. 3 a</p> <p>4 c) Leistungen für Praxis und Gesellschaft werden nicht berücksichtigt (abgewertet) (populärwissenschaftliche. Publikationen, Vorträge, ...)</p> <p>4 d) Kooperation wird durch Konkurrenzdruck erschwert</p>	<p>8/22 Leistungen in der Praxis können nicht publiziert werden, etablierte Kriterien haben zu hohen Stellenwert bei Besetzung von Professuren</p> <p>1/22</p>
<p>5) zu häufige Evaluationen</p>	<p>10/22 reduzieren Qualität, Orientierung am Output</p>
<p>Drittmittelvolumen (Ergänzungen der Interviewpartner)</p>	<p>3/22 drängt die Forschung in die privatwirtschaftliche Verwertung oder die Grundlagenforschung (DFG-Mittel = „qualifizierte“ Drittmittel), Ökolandbau-Forschung mit gesellschaftlichem Nutzen aber ohne vermarktbare Produkte wird benachteiligt</p>

4 Berücksichtigung von Leistungen der Forschung für Praxis und Gesellschaft in der Forschungsevaluierung

In diesem Kapitel werden die Ergebnisse aus der Literaturanalyse zu bestehenden Konzepten für die Evaluierung angewandter, inter- und transdisziplinärer Forschung ebenso dargestellt wie die Ergebnisse aus den Interviews und dem Workshop des Projektes. Kapitel 4.1 stellt überblicksartig die als relevant erachtete Literatur zum Thema dar. In Kapitel 4.2 werden dagegen die detaillierten Ergebnisse aus den Interviews und dem Workshop präsentiert und auf Basis einer intensiven Auswertung der Literatur diskutiert.

4.1 Stand des Wissens zur Berücksichtigung von Leistungen für Praxis und Gesellschaft in der Forschungsförderung und Forschungsevaluierung

4.1.1 Ziele öffentlich geförderter Forschungsprogramme der Agrar- und Nachhaltigkeitsforschung

In den Programmzielen der öffentlich geförderten Agrar- und Nachhaltigkeitsforschung wird in vielfältiger Weise ein Bezug zu Problemen von Praxis und Gesellschaft hergestellt (vgl. Tabelle 3). Forschung soll einen Beitrag zu einer nachhaltigen Entwicklung leisten. Darüber hinaus wird ein Wissenstransfer in die Praxis bis hin zur transdisziplinären Forschung grundsätzlich verlangt. In einigen Programmen wird auch explizit eine interdisziplinäre Herangehensweise begrüßt. Mit der etablierten Forschungsevaluierung können diese Zielsetzungen nicht oder nur eingeschränkt überprüft werden.

4.1.2 Anforderungen im Zuge der Beantragung von Drittmitteln in der EU und USA

Kriterien für die Bewertung von Anträgen sind bereits jetzt vielfach weiter gefasst, als die klassischen Evaluierungskriterien. Das 7. EU Rahmenprogramm gibt für die Antragsbewertung in den Programmteilen „Zusammenarbeit“ und „Kapazitäten“ vier Kriterien vor: (a) Relevanz für die Erreichung der Programmziele, (b) Wissenschaftliche/technologische Exzellenz, (c) potenzieller "Impact" durch Entwicklung, Verbreitung oder Verwertung der Projektergebnisse, (d) Qualität und Effizienz von Projektimplementierung und –management (EC 2006).

Forschungsanträge an die US-amerikanische National Founding Society (NSF) müssen sowohl den "Intellectual Merit" als auch die "Broader Impacts" beschreiben, wobei sich letzteres auf Bildung, breitere Partizipation unterrepräsentierter Gruppen, Verbesserung der Forschungsinfrastruktur, Verbreitung von Ergebnissen und den gesellschaftlichen Nutzen bezieht. (NSF 2007). Die Beobachtung des seit 2001 eingeführten Kriteriums "Broader Impacts" zeigte aber auch deutlich, wie fremd für Teile der Scientific Community eine engere Interaktion zwischen Wissenschaft und Gesellschaft ist (Frodeman und Holbrook 2007).

4.1.3 Ausgewählte Leitbilder der deutschen Bundesministerien

Auch durch die deutschen Ministerien wird vielfach eine integrierte Zielsetzung für die öffentlich geförderte Forschung formuliert, die sowohl der wissenschaftlichen Qualität als auch dem gesellschaftlichen Nutzen verpflichtet ist

Tabelle 3: Zielsetzungen ausgewählter öffentlicher Förderprogramme im Bereich Agrar- und Nachhaltigkeitsforschung

Forschungsprogramm	Bezug zu Problemen von Praxis und Gesellschaft bis hin zur Transdisziplinarität	Wissenstransfer in Praxis und Gesellschaft	Interdisziplinarität	Beiträge zur nachhaltigen Entwicklung
BMBF: FONA Forschung für nachhaltige Entwicklungen	Transdisziplinarität = alle notwendigen Partner des Innovationsprozesses sowie alle relevanten gesellschaftlichen Gruppen werden frühzeitig eingebunden; neues Wissen muss auch für die Praxis relevant sein. (BMBF 2009:2)	Transfer von Forschungsergebnissen in die Praxis u. a. über Kompetenzzentren (BMBF 2009:37, 56); Stärkung von Infrastrukturen der transdisziplinären Nachhaltigkeitsforschung; Sichtbarkeit der Ergebnisse durch Publikations- und Transfermaßnahmen unterstützt. (BMBF 2009:40)	trans- und interdisziplinäre Forschungskonzepte ausbauen; Untersuchungen zu gesellschaftlicher und kultureller Entwicklung (BMBF 2009:8); interdisziplinäre Verbundprojekte; problem- oder systemorientiert (BMBF 2009:20); strukturbildende Maßnahmen an interdisziplinären Schnittstellen (BMBF 2009:20); ressortübergreifende Zusammenarbeit (BMBF 2009:38)	Forschung als Vorsorge für die Zukunft (BMBF 2009:2); integrierter konzeptioneller Forschungsansatz, der wirtschaftliche, ökologische und soziale Aspekte gleichzeitig umfasst (BMBF 2009:5); zentrale, dringende Themen: Klima, Energie und Ressourcen, Systemansatz entscheidend (BMBF 2009:8); Nachhaltige Landnutzung interdisziplinär im gesamtgesellschaftlichen Kontext betrachtet (BMBF 2009:44)
BfN: Bundesprogramm Biologische Vielfalt	Multiplikatorwirkung; Kooperation unterschiedlicher Akteure bei der Umsetzung der Ziele (BMU 2011)	Akzeptanz bildende Maßnahmen der Information und Kommunikation sollen dazu beitragen, das gesellschaftliche Bewusstsein für die biologische Vielfalt zu stärken (BMU 2011)	(nicht explizit erwähnt)	Rückgang der biologischen Vielfalt stoppen bzw. Trend umkehren; Schutz und nachhaltige Nutzung der biologischen Vielfalt (BMU 2011)
BMELV und BLE: Innovationsförderung	Ziel: Innovationen in der deutschen Agrar- und Ernährungswirtschaft sowie im Verbraucherschutz (BMELV 2009)	Wissenstransfer (BMELV 2009)	(nicht explizit erwähnt)	Schonung von Ressourcen; Züchtung robuster und leistungsfähiger Kulturpflanzen; neue Haltungsverfahren zur Verbesserung des Tierschutzes (BMELV 2009)
BMELV und BLE: BÖL, seit 2011 BÖLN (Bundesprogramm Ökologischer Landbau und andere Formen nachhaltiger Landwirtschaft)	Beitrag zur Lösung von Praxisproblemen (BMELV 2010); Einbindung der Praxis bereits bei der Projektplanung (BÖLN 2011)	breiter, systematischer Wissens- und Technologietransfer (www.forschung.oekolandbau.de) (BMELV 2010)	drei interdisziplinäre Förderschwerpunkte (Milchvieh, Sauen/Ferkel, Bodenfruchtbarkeit) (BÖLN 2011)	ökologisch bewirtschaftete Fläche erhöhen; Wissens- und Erfahrungslücken schließen; Wettbewerbsfähigkeit des Ökologischen Landbaus erhöhen; Ökolandbau als Leitbild für die Landwirtschaft: stärkt und bewahrt langfristig die Lebensgrundlage des Menschen (BMELV 2010)
BMELV und BLE: Fachagentur nachwachsende Rohstoffe (FNR)	produktions- und verwendungsorientierte anwendungsbezogene Forschungs-, Entwicklungs- und Demonstrationsvorhaben (BMELV 2008: 5)	Veröffentlichungen für vielfältige Zielgruppen, Messen und Ausstellungen; öffentliche Zugänglichkeit von Forschungsergebnissen; Beratung und Öffentlichkeitsarbeit (BMELV 2008: 23)	Erhöhung der Forschungsmittel für inter- und transdisziplinäre Forschungsprojekte (BMELV 2008: 22); wissenschaftlicher Erfahrungsaustausch im nationalen und internationalen Rahmen (BMELV 2008: 20)	Lösung von wirtschafts-, umwelt- und gesellschaftsrelevanten Problemen; Kreislaufwirtschaft; nachhaltige Wirtschaftsformen; Schonung fossiler Ressourcen; Stärkung des ländlichen Raums; Erhaltung der biologischen Vielfalt (BMELV 2008:4f)

4.1.4 Evaluierungskonzepte für die inter- und transdisziplinäre Forschung

Für die inter- und transdisziplinäre (Nachhaltigkeits-)Forschung wurden in vielen Ländern Evaluierungskonzepte erarbeitet, die in Summe ein breites Spektrum an Kriterien, Instrumenten und Vorgehensweisen umfassen. Praxisorientierte Forschungsleistungen werden aber auch in einzelnen Konzepten zur Agrarforschung berücksichtigt. Auch die Ressortforschung wird regelmäßig evaluiert (vgl. Wissenschaftsrat 2007). Um ihre Kernaufgabe, die wissenschaftliche Politikberatung, angemessen zu berücksichtigen geht sie dabei in ihren Qualitäts- und Evaluierungskriterien über den wissenschaftlichen Impact hinaus. Exemplarisch wurde das Konzept für die Ressortforschung im BMU von Böcher und Krott (2010) detailliert ausgewertet.

Die in der Literaturstudie ausgewerteten Konzepte sind in Tabelle 4 aufgeführt. Sie beziehen sich auf unterschiedliche Evaluierungsgegenstände (Wer/Was wird evaluiert?) und –zeitpunkte (In welcher Forschungsphase wird evaluiert?).

Tabelle 4: Ausgewertete Literatur zur Evaluierung inter- und transdisziplinärer (Nachhaltigkeits-)forschung sowie praxisorientierter Agrarforschung.

Reference	object of research evaluation		Application in research stage
1 Aenis und Nagel 2003	trans	Projekt	on-going, ex post
2 Albrecht et al. 2008, 43f	inter /trans	Projekt	ex post
3 Alföldi et al. 2007	agrар	Projekt	ex ante
4 Begusch-Pfefferkorn 2006	inter / trans	Programm	ex post
5 Bergmann et al. 2005	trans	Projekt	on-going
6 Blackstock Kelly und Horsey 2007	trans	Projekt	ex post
7 Boix Mansilla and Gardner 2003	inter	Institution/Forschungsgruppe	ex post
8 Böcher und Krott 2010	Ressortforschung	Institution/Forschungsgruppe	ex post
9 Defila et al. 2008, 72f	inter / trans	Forschungsgruppe	on-going
10 Defilia und Di Giuilo 1999	trans	Projekt	mehrfach
11 ERIC 2010	Forschung generell	Institution/ Forschungsgruppe	ex post
12 Formas 2009	agrар	Projekt	ex post
13 Glaeser 2007	inter / trans	-	ex ante
14 Krott 2004	trans	-	-
15 Lyall et al. 2004	agrар (Biologie)	Institution	ex ante
16 Pedersen et al. 2009	agrар	Projekt	ex post
17 Schiller et al. 2005	inter/(trans)	Projekt/ Forschungsgruppe ⁴	ex post /(ex ante)
18 Schübel 2007	trans	Projekt	on-going, ex post
19 Smoliner 2007	trans	Programm	ex post
20 Spaapen et al. 2007	inter / trans	Forschungsgruppe	ex post
21 Stokols et al. 2003	inter /(trans)	Institution/ Forschungsgruppe	on-going /mehrfach
22 Stoll-Kleemann 2007	inter / trans	Projekt	-
23 Walter et al. 2007	trans	Projekt, Programm	on-going, ex post
Die Nummerierung wird verwendet, um die Interviewergebnisse mit den Inhalten der Literatur zu vergleichen. Evaluierungsgegenstände sind sowohl die inter- und transdisziplinäre Forschung sowie die Agrarforschung als Projekte, Programme, Forschungsgruppen und Institutionen.			

⁴ Einzelne Institute einer Universität

Der Schwerpunkt der ausgewerteten Konzepte liegt auf der Evaluierung von Projekten und der ex post Perspektive. Die Ziele, Kriterien und Evaluierungsinstrumente der ausgewerteten Konzepte werden gemeinsam mit den Interview- und Workshopergebnissen in Kapitel 4.2 - 4.5 dargestellt.

Weitere Quellen für Kriterien, die Leistungen für die Gesellschaft berücksichtigen sind das SIAMPI Projekt (SIAMPI 2010), welches die Konzeption von ERiC 2010 weiterführt, jedoch noch keine abschließenden Ergebnisse veröffentlicht hat. Des Weiteren wurde ein societal impact assessment in Verbindung mit wissenschaftlichen Publikationen von der Medizinischen Universität Wien dem wissenschaftlichen Publikum zur Begutachtung vorgestellt, welcher ein Bewertungsraster für Transferleistungen und Impact mit einer Indexbildung beinhaltet (Niederkrötenhaller, Dorner und Maier 2011).

Hinsichtlich des Zwecks einer Evaluierung, kann zwischen miteinander verbundenen Funktionen differenziert werden: Erkenntnis, Lernen, Kontrolle, Legitimierung (Stockmann 2004). Hauptfokus der ausgewerteten Evaluierungskonzepte sind Lernprozesse zur Qualitätsverbesserung (Aennis and Nagel 2003, Boix Mansilla und Gardner 2003, Stokols et al. 2003, Bergmann et al. 2005, Blackstock, Kelly und Horsey 2007, Schübel 2007, Walter et al. 2007). Diese Konzepte beinhalten detaillierte Kriterien für den Forschungsprozess und sind meist sehr spezifisch auf den idealtypischen Ablauf transdisziplinärer Projekte zugeschnitten (z. B. Bergmann et al. 2005, 17). Aber auch die eher kontrollierende/legitimierende Zwecksetzung von Evaluierung wird aufgeführt, z. B. gegenüber dem Geldgeber (Aennis and Nagel 2003, Stokols et al. 2003, Begusch-Pfefferkorn 2006, Smoliner 2007), durch einen Soll-Ist-Vergleich (Defila und Di Giulio 1999) oder ein Benchmarking zwischen Institutionen bzw. Forschungsgruppen (Spaapen, Dijstelbloem und Wamelink 2007), sowie die grundsätzliche Verdeutlichung des gesellschaftlichen (Aennis und Nagel 2003, Stokols et al. 2003) oder praktischen Nutzens (Walter et al. 2007).

Trotz erheblicher Unterschiede in den Zielsetzungen sowie im Detaillierungsgrad und der Anzahl der Kriterien (einige Ansätze beschreiben nur Methoden), liegen deutliche Übereinstimmungen dazu vor, welche praxisorientierten und interdisziplinären Leistungen der Forschung berücksichtigt werden (s. Tabelle 6 und Tabelle 7). Im Sinne eines Gesamtkonzeptes wird in den Vorschlägen die wissenschaftliche Qualität und Nachwuchsförderung, berücksichtigt, die hier jedoch nicht im Fokus stehen, da eine Ergänzung der etablierten wissenschaftlichen Evaluierung angestrebt wird.

Die bestehenden Evaluierungskonzepte unterstreichen: „Quality is a relative concept, driven by variability of goals and criteria“ (Klein 2006). „Feldspezifische“ Zielsetzungen, in Verbindung mit anerkannten Standards für die Evaluierung (DeGEval 2008), stellen somit einen geeigneten Rahmen für die Entwicklung von Evaluierungskonzepten dar.

4.2 Ergebnisse der Befragung und Gruppendiskussion diskutiert mit einschlägiger Literatur.

Die Ergebnisse umfassen die Darstellung der Interviewergebnisse und Gruppendiskussion, ergänzt durch einschlägige Literatur zu:

- a) Anforderungen und Ziele für die Evaluierung praxisorientierter Ökolandbauforschung inkl. einer Bewertung der Kriterienbereiche aus dem Interviewleitfaden beschreiben den Rahmen für das Evaluierungskonzept
- b) die Kriterien für die Bereiche Praxisorientierung, Interdisziplinarität, Nachhaltigkeit, Gesellschaftliche Relevanz, Systemorientierung, Vielfalt von Themen und Methoden, Innovation

und Originalität sowie Langfristigkeit der Forschung beschreiben die inhaltliche Ausgestaltung der Evaluierung

- c) die Analyse geeigneter Evaluierungsinstrumente skizziert die Methodik des Evaluierungskonzeptes

4.3 Anforderungen und Ziele für die Evaluierung praxisorientierter Ökolandbauforschung

4.3.1 Individuelle Anpassung der Evaluierung

Als wichtigste Anforderungen an die Evaluierung formulierten die Interviewpartner die Notwendigkeit eines breiten, anpassbaren Kriteriensets (11/22 + 7/10 WS). Dazu gehört, dass die Evaluation gleichermaßen den Output (z. B. Publikationen und Produkte), den Prozess (z. B. Praxis-Kontakte, praxisorientierte Arbeitsweise) die Relevanz der Fragestellung und Ergebnisse sowie die tatsächliche Wirkung in der Praxis umfasst (11/22). Für die Evaluierung sollten sowohl qualitative als auch quantitative Kriterien verwendet werden (7/22). Ein breites Kriterienset ist notwendig, um *erstens* die Individualität der Forschung zu berücksichtigen, und *zweitens* verschiedenen Charakteristika und Anforderungen der jeweiligen Zielgruppen der Forschung (Nutzer, Betroffene) gerecht zu werden. *Drittens* unterscheiden sich auch die Anforderungen an eine Evaluierung hinsichtlich der Evaluierungsgegenstände (Wissenschaftler, Projekte, Institutionen, Programme) selbst, dem Kontext der jeweiligen Evaluierungsgegenstände und den Zielsetzungen der Evaluierung. Die genannten Aspekte erfordern eine Anpassung des grundsätzlich breiten Kriteriensets an jede spezifische Evaluierung. Diese Anpassung umfasst u. a. die Auswahl der Kriterien (4/22 + 7/10 WS), als auch die Auswahl geeigneter Akteure aus Praxis und Gesellschaft (8/21), die für die Evaluierung relevant sind. Bei der Interpretation der Evaluierungsergebnisse (3/10 WS) sind u.a. Leistungen in Relation zu den vorhandenen Ressourcen zu bewerten (2/22). Denn eine bessere Ausstattung mit Finanzmitteln und Personal verbessert die Möglichkeiten zur Erbringung von Leistungen für Praxis und Gesellschaft. Eine Übersicht zu diesen Anforderungen ist in Abbildung 3 dargestellt.

Der individuellen Anpassung der Evaluierung wird von den Befragten insbesondere deshalb hohe Bedeutung beigemessen, weil Forschung als originärer, individueller Prozess verstanden wird, der auf vielfältige Weise Wert und Wirkung entfalten kann. Gerade die Interviewergebnisse zu den Zielgruppen zeigen, dass Ökolandbauforschung gefordert ist, Lösungsbeiträge sowohl für eng umrissene Zielgruppen (z. B. Bauern und Bäuerinnen einer bestimmten Betriebsform) als auch für regionale und globale Herausforderungen einer nachhaltigen Entwicklung zu leisten.

Auch in der Literatur gibt es Konzepte, in denen die Bedeutung des Kontextbezuges in der Evaluierung besonders hervorgehoben und umgesetzt wird z. B. bei Spaapen, Dijstelbloem und Wamelink (2007), Blackstock, Kelly und Horsey (2007). Ein anderer Evaluierungsansatz ist die Konzentration auf die Voraussetzungen zum Projektbeginn, statt auf die Leistungen zum Projektende, da die langfristigen Wirkungen der Forschung als unzureichend erfassbar eingeschätzt werden (Aennis und Nagel 2003).

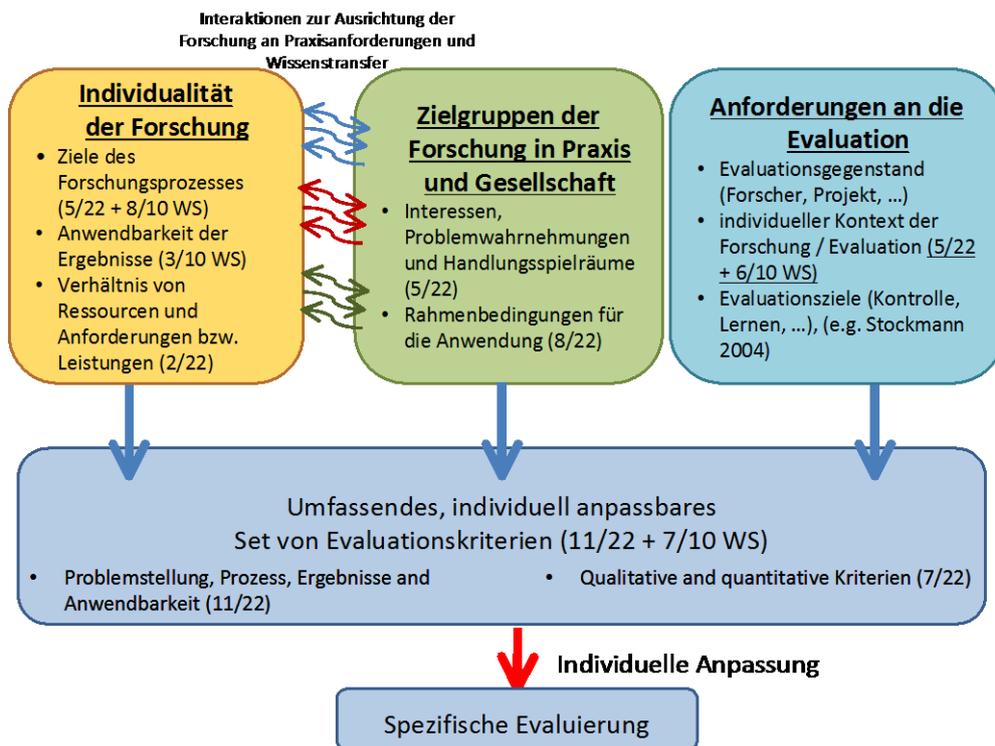


Abbildung 3: Individuelle Anpassung der Evaluierung

4.3.2 Bewertung verschiedener Kriterienbereiche

Die grundsätzlichen Einschätzungen und die Kernaussagen zu den im Interviewleitfaden aufgelisteten Kriterienbereichen Tabelle 5 geben Aufschluss darüber, welchen Stellenwert diese in einem Evaluierungskonzept einnehmen könnten. In den Interviews wurde insbesondere den Leistungen der Forschung für die Praxis (Praxisimpact) und der Interdisziplinarität ein hoher Stellenwert beigemessen. Für diese Bereiche wurde auch eine Vielzahl konkreter Kriterien genannt, die auch in den Evaluierungskonzepten anderer Autoren relevant sind (vgl. Tabelle 6 und Tabelle 7).

Die weiteren im Interviewleitfaden erfragten Kriterienbereiche wurden grundsätzlich befürwortet, aber, wie aus den Kernaussagen ersichtlich, wesentlich kontroverser diskutiert. Sowohl die Interviewergebnisse, als auch die bestehenden Evaluierungskonzepte nennen Kriterien für diese Bereiche, die mehrheitlich qualitativ und in einigen Fällen auch weniger genau spezifiziert sind (vgl. Tabelle 8).

Fragen wie Nachhaltigkeit Vielfalt von Themen oder gesellschaftliche Relevanz müssen, auch wenn ihre Bewertung schwieriger ist, in der Forschungsevaluierung eine Rolle spielen, da sie die Leitbilder der öffentlichen Forschungsförderung, (vgl. auch Tabelle 3) als auch dem Selbstverständnis verschiedener Forschungsrichtungen fundamental berühren. Beispielsweise der Bezug zu lebensweltlichen Problemen als tragendes Merkmal transdisziplinärer (Nachhaltigkeits-)forschung (vgl. Bergmann et al. 2005, 15, Pohl und Hirsch Hadorn 2006, 22).

Für die Agrarforschung wird im IAASTD-Bericht die Berücksichtigung von Gerechtigkeit, ökologischer und sozialer Nachhaltigkeit, (unter Einbindung lokaler, betroffener Akteure) in der Forschungsevaluierung gefordert (IAASTD 2009, 217). Für den Ökologischen Landbau sind dessen spezifische Anforderungen an die Wissenschaft zu berücksichtigen. Diese werden sowohl vornehmlich aus Sicht der Wissenschaft wie z.B. Lindenthal, Vogl und Heß (1996), Schmid et al. (2009) als auch, in weniger umfangreichem Rahmen, aus Sicht der Praxis (Pioniere des Ökolandbaus) z.B. durch Aeberhard und Rist (2009) dargestellt.

Darüber hinaus wird von umweltpolitisch und sozial engagierten Initiativen innerhalb und außerhalb der Wissenschaft zunehmend eine verstärkte Ausrichtung an Problemlagen und einer demokratischen Kontrolle der Vergabe von öffentlichen Forschungsgeldern gefordert. (z. B. VDW 2010).

Tabelle 5: Kriterienbereiche aus dem Interviewleitfaden - Bewertung und wichtige inhaltliche Aussagen

Bewertung der Kriterienbereiche	wichtige inhaltliche Aussagen zum Kriterienbereich
Praxisorientierte Leistungen sehr wichtig (21/22) ablehnend (1/22) (stattdessen Peer-Review Journals für praxisorientierte Forschung etablieren)	<ul style="list-style-type: none"> - Praxis muss als Korrelat zu Theorie unbedingt in der Forschung integriert sein, aber Wirkungen in der Praxis sind nicht das einzige Kriterium für Forschungsevaluierung - Bewertung erfordert Einbindung der Praxis in die Evaluierung
Interdisziplinarität sehr wichtig (19/22) nicht notwendig (3/22)	<ul style="list-style-type: none"> - trägt zu problemgerechteren Antworten bei, ist innovationsstiftend, schafft Vielfalt von Themen und Methoden - Definition und Abgrenzung von Disziplinen ist problematisch
Gesellschaftliche Brisanz und Relevanz sehr wichtig, aber schwierig bewertbar (13/22) skeptisch (4/22)	<ul style="list-style-type: none"> - Bezieht sich auf die Identifikation relevanter Probleme und die Bewertung von Lösungsbeiträgen und deren gesellschaftliche Auswirkungen - Mitbestimmungsrecht der Gesellschaft - Verantwortung der Forschung gegenüber d. Gesellschaft und öffentlichen Gütern – gleichzeitig Konflikt mit Freiheit der Forschung - schwer objektivierbar - Gesellschaft ist vielfältig und wandelbar – wer wird berücksichtigt?
Vielfalt von Themen und Methoden grundsätzlich befürwortet: Themenvielfalt (16/22) Methodenvielfalt (6/22)	<ul style="list-style-type: none"> - eng verbunden mit Inter- und Transdisziplinarität - Vielfalt ist notwendig für Innovation und Flexibilität für zukünftige Problemlagen - Vielfalt und Spezialisierung sollten integriert werden
Originalität wichtig (14/22) wichtig unter Vorbehalt (6/22) ablehnend (2/22)	<ul style="list-style-type: none"> - schwer objektivierbar, oft erst rückblickend erfassbar - Originalität soll sich eingliedern in Vielfalt, Systembezug, Bezug auf Problemlagen von Praxis und Gesellschaft
Normative Kriterien (z. B. Systembezug, Nachhaltigkeit, Prinzipien Ökolandbau) wichtig (11/22) skeptisch (4/22) ablehnend (2/22)	<ul style="list-style-type: none"> - Intensive Diskussion darum, welche Kriterien normativ sind und ob Wissenschaften u. Forschungsprogramme wertfrei sind, sein sollen bzw. überhaupt sein können⁵ - Verantwortung ggü. Umwelt und Gesellschaft vs. Einschränkung der Freiheit von Lehre und Forschung - bestimmte Werte sind Grundbausteine des Ökolandbaus schwer objektiv zu bewerten
Systembezug wichtig (10/22) aber Symptomorientierung nicht ausschließen (2/22)	<ul style="list-style-type: none"> - aufwändig zu erheben - Systemgrenzen schwer zu definieren
Nachhaltigkeit wichtig (9/22)	<ul style="list-style-type: none"> - wurde von den Interviewpartnern als eigener Kriterienbereich vorgeschlagen
Langfristigkeit wichtig (7/22)	<ul style="list-style-type: none"> - betrifft sowohl längere Zeiträume zur Lösungsfindung, als auch langfristige Problemlösungen und langfristige Auswirkungen - Forschungsförderung hat einen größeren Einfluss auf die Ermöglichung langfristiger Forschung als angemessene Evaluierung

⁵ Diesen Fragen liegen unterschiedliche Paradigmen bezüglich Objektivität und Wertfreiheit der Wissenschaft zugrunde (z. B. Positivismus und Konstruktivismus) (vgl. Lee 2004 in Stockmann 2004).

4.4 Kriterienvorschläge in Interviews, Workshop und Literatur

Tabelle 6 und Tabelle 7 zeigen die vorgeschlagenen Einzelkriterien für die Bewertung von Praxisbezug und Interdisziplinarität wofür sehr umfangreiche Kriterien sowohl in den Interviewergebnissen, als auch in der Literatur vorgeschlagen werden. Tabelle 8 zeigt die (weniger zahlreichen) Kriterien für die Bewertung von Nachhaltigkeit, gesellschaftlicher Relevanz, Systemorientierung, Vielfalt von Themen und Methoden, Innovation und Originalität und Langfristigkeit der Forschung aus Interviews und Workshop sowie Literatur.

Die Verweise auf ähnlichen Kriterien in der Literatur erfolgen mit Nummern aus Tabelle 4. Aufgeführt werden auch als ergänzenswert eingeschätzte oder genauer spezifizierte Kriterien aus der Literatur. Die Kriterien können, je nachdem wie detailliert oder aggregiert sie dargestellt werden, gleichermaßen für verschiedene Evaluierungsgegenstände (Wissenschaftler, Projekte, Institutionen, Forschungsprogramme) verwendet werden. Einige Kriterien sind lediglich für Programme (P) oder Institutionen (I) anwendbar und entsprechend gekennzeichnet.

Die Kriterien werden in vier Gruppen (A-C) unterteilt, die sich in ihrer Komplexität und den für ihre Anwendung notwendigen Evaluierungsinstrumente unterscheiden. Mit dieser Einteilung erfolgt auch eine Abschätzung der Anwendbarkeit der Kriterien und des damit verbundenen Aufwandes. Inwieweit und wie einzelne Kriterien umgesetzt werden können, hängt aber auch von den Einstellungen und Zielen der Beteiligten einer Evaluation (Auftraggeber, Evaluatoren, Evaluierete) und den vorhandenen Ressourcen für die Evaluierung ab und kann damit im Detail erst in einem konkreten Evaluierungsvorhaben definiert werden.

- A bei rein quantitativer Bewertung mit geringem Aufwand erfassbare Kriterien innerhalb der Wissenschaft gut zu dokumentieren, z. B. durch Erweiterung der Forschungsdatenbanken an Universitäten
- B Kriterien, die bei einer kombinierten quantitativen und qualitativen Bewertung innerhalb der Wissenschaft mit mäßigem Aufwand bewertet werden können, (erfordern mehr Informationen durch den Wissenschaftler z. B. Dokumentation der Wissenschaftler in einem Datenbanksystem oder Interviews/Gruppendiskussionen mit Wissenschaftlern, Auswertung von Endberichten etc.)
- C Für die Bewertung ist die Einbindung von nicht-wissenschaftlichen Akteuren erforderlich/sinnvoll, womit je nach Methode ein mäßiger bis höherer Aufwand notwendig ist (z. B. Interviews, Gremien, Praktikerbeiräte, Gruppendiskussionen, Gutachter aus der Praxis).
- D Vervollständigung und Konkretisierung der Kriterien ist erforderlich (bspw. im Vorfeld der Evaluierung oder durch weitere Forschungsarbeiten).

Tabelle 6: Kriterien für die Bewertung praxisorientierter Leistungen von Wissenschaftlern, Projekten, Institutionen und Programmen.⁶

Bereich (Anzahl Personen mit Vorschlägen)	Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwendbarkeit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
Praxisbezogene Kompetenzen und Weiterbildung des Wissenschaftlers (2/22 + 1 WS)	Kompetenzen und Erfahrungen in praxisorientierter Forschung (1)	A, B	3, 5, 14, 21
	Weiterbildung zu Kommunikationskompetenzen etc. (2)	A, B	Besuch praxisorientierter Weiterbildungsveranstaltungen (4)
Veröffentlichungen und Produkte für die Praxis 13/22 + 5/10 WS (relevante Leistungen)	Artikel in Fachzeitschriften (9) Beratungsempfehlungen (Q 14/15) Fachbücher Beiträge in Zeitungen etc. (2) Praxisinformationen in digitalen Medien (1) andere Publikationen und Produkte (1)	A, B	4, 11, 17, 22 generell Publikationen und Produkte für die Praxis (2, 3, 5, 8, 10, 15, 20)
	Gutachten/Anhörungen (2)	A, B	
	Produkte (z. B. IT-Tools, Patente, Maschinen, Konzepte) (1)	A, B	3, 5, 10, 15, 20
Praxisorientierte Arbeitsweise/Praxiskontakte 14/22 + 5/10 WS (relevante Leistungen)	Praxiskontakte (Q 16/16) mit konkreten Beispielen (8), u. a. Workshops, Feldtage, Arbeitskreise/Gremien, Kooperationen, Umsetzungsprojekte, Kooperationen mit Praxis oder Praxis-Institutionen. Vorträge für die Praxis (9) Vielfalt der Kontakte muss über qualitative Aspekte spezifiziert werden.	A, B	allg. Praxiskontakte (4, 11, 13, 20, 22), Vorträge (4, 5, 10, 11, 17, 22) Transferleistungen (2) Anzahl Praxispartner (3, 4, 11, 20) repräsentative Einbindung verschiedener Sichtweisen (6) I: Beratung für Bürger (8) I: Kommunikation mit „End-Nutzern“ (15)
	Netzwerkbildung mit der Praxis (Q 14/16) (national befürwortet, international eher skeptisch)	B, C	4, 5, 11, 23
	„capacity building“ (dauerhafte Strukturen, verbesserte Rahmenbedingungen für die Praxis) (3)	B, C,	6
	Kofinanzierung aus der Praxis (2 positiv, 1 negativ)	A, B	4, 11, 17, 20
Potenzielle Praxisrelevanz	Lösungsbeitrag zu Praxisproblemen (8) Potenzieller Nutzen (8) und Risiken (3) für Zielgruppen, Gesellschaft und Umwelt Größe und Charakteristika der Zielgruppe (5 + 2 WS) Wichtigkeit (1), Umsetzbarkeit (2), Nützlichkeit (1)		3, 5, 6, 8, 10, 12, 14, 16, 20, 22, 23 Interesse von Stakeholdern (11) Übergabestrategie/Unterstützung für Anwendung (5, 10) Umsetzungsstrategie/ Kenntnis der relevanten Akteure (8), Machbarkeit (8, 17), I: generelle Relevanz einer Institution und ihrer Forschungsthemen 15)
Qualitätskriterien für Publikationen und Produkte (P), Kontakte mit Praxisakteuren (K) und Relevanz für die Praxis (R)	P Thematische Vielfalt oder Publikation eines kontinuierlichen Wissensfortschritts (5)	B	
	P, K Einladungen aus der Praxis (4 + 1 WS) (nur in einigen Fällen relevant)	A, B	
	P, K Aufbereitung von Ergebnissen und ausgewählte Medien(8 + 2 WS) sind Zielgruppengerecht	C	5, 8, 10, 12, 15, 16
	K Strukturen und Methoden für Interaktion und Rückkopplung (9)	B, C	5, 10, 14, 22, 23, allg. Integration (2) Neue Konzepte, Methoden, Theorien zu Transdisziplinärer Forschung (5, 21), Transparenz projektintern und extern (6)
	K Qualität der Kommunikation/Kooperation (8) Lernprozesse / Synthese zwischen wissenschaftlichem und praktischem Wissen (5)	B, C	5, 6, 7, 10, 14, 15, 22, 23

⁶ Die Kriterien sind prinzipiell für alle Evaluierungsgegenstände geeignet, müssen aber jeweils wahlweise detailliert oder stärker aggregiert erfasst werden. Einige Kriterien sind mit P oder I gekennzeichnet, da sie sich nur auf Programme oder Institutionen beziehen.

Bereich (Anzahl Personen mit Vorschlägen)		Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwendbarkeit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
	K, R	Berücksichtigung von Problemen (6) und Erfahrungswissen (1 + 10 WS) der Praxis	B, C	5, 7, 10, 14, 15, 22, 23 Ausreichend fachliche Breite zur Problemlösung insbes. bezüglich Problemursachen, naturwiss./ökon./rechtl. Zusammenhänge (17) Berücksichtigung der Machbarkeit (8, 17), Berücksichtigung von lokalem Wissen (6), Übersetzung von Praxisproblemen in wiss. Fragen (5, 8), allg. Problemorientierung (2)
Nachbeobachtung: Anwendung, Wirkung und Lösungsbeitrag in der Praxis 21/22 + 2/10WS		Resonanz in Praxis und Gesellschaft z. B. - Verbreitung der Ergebnisse in der Praxis (Erreichen relevanter Zielgruppen, Multiplikatoren etc.) (4 + 2 WS) - Intensität der Diskussion über ein Thema - Nachfrage nach Publikationen und Produkten - Folgeprojekte mit der Praxis (1)	C	Nationale und internationale Auszeichnungen der Praxis, Bildung oder Politik (4, 17), Berufliche Mobilität in die Praxis (7, 17, 20)
		Anwendung (oder modifizierte Anwendung) von Forschungsergebnissen, Konzepten, Produkten etc. in der Praxis (und Gesellschaft) (20) - Umfang der Nutzung: z.B. Anzahl der Nutzer oder betroffene Fläche - Größe und Charakteristika der Zielgruppe (5 + 2 WS) - Sichtbarkeit in Agrarstatistiken/ Verkaufszahlen etc. (3) - Anzahl Lenkungseffekte auf das Förderwesen/Sachpolitik (Q 12/16)	B, C	1, 5, 7, 8,10, 11, 15, 21, 22, 23 Verkaufszahlen etc. (11, 14) Lenkungseffekte (4,11)
		Auswirkungen und Lösungsbeitrag in der Praxis (7) z. B. - Nutzen für Zielgruppe, Umwelt und Gesellschaft - Vor- und Nachteile der Anwendung/Wirkung - Forschungsdefizite und Lösung von Praxisproblemen –(Projektübergreifend erhoben) (1)	B, C	5, 11, 14, 6 Wurde ein Lösungsbeitrag erarbeitet? Wiss. Bewertung (5, 10)

Tabelle 7: Kriterien für die Bewertung der Interdisziplinarität.⁷

Bereich (Anzahl Personen mit Vorschlägen)		Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwendbarkeit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
Interdisziplinäre Kompetenzen und Weiterbildung des Wissenschaftlers (2/22)		Kompetenz, Erfahrung und bisherige Zusammenarbeit der Projektpartner (1),	A, B	3, 5, 14, 21,
		Kurse für interdisziplinäre Arbeitsmethoden (1)	A, B	
Interdisziplinäre Veröffentlichungen und Produkte (6/22)		Interdisziplinäre Publikationen/Bücher/Endberichte (5)	A	2,17
		Autorengemeinschaften aus versch. Disziplinen (5)	A	4,17,

⁷ Siehe Fußnote 6

Bereich (Anzahl Personen mit Vorschlägen)	Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwendbarkeit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
Disziplinäre Vielfalt (10/22 + 1/10 WS)	Allg. unterschiedliche Themengebiete/Disziplinen (9)	A	4, 21
	Anzahl Disziplinen (5 befürwortend, 2 ablehnend)	A, B	Befürwortend (4) ablehnend (21)
	Beteiligung der notwendigen Disziplinen (5)	B (D)	5, 17, sich ergänzende Fachkompetenz (3)
	Entfernung zwischen Disziplinen (4)	B (D)	
Interdisziplinäre Arbeitsweise (14/22)	Forschungsmethodik fördert Synthese zw. Disziplinen (5) Strukturen und Methoden für den kooperativen Arbeitsprozess (3)	B	5,7,10, 14, 17, allg. interdisziplinäre Integration (2)
	P/I: Angemessenes Verhältnis von Ressourcen und Anforderungen (7) (auch relevant für praxisorientiertes Arbeiten)	A	10, 21, Institutionelle Unterstützung (17)
	Intensität des Austauschs/ der Vernetzung (8) (z.B. gemeinsame Problemdefinition, Aufgabenverteilung, Methodikentwicklung, ggs. Lernprozesse zu Begriffen, Themen, Methoden und Wertorientierungen der anderen Disziplinen)	B, (D)	3, 4, 5, 7, 10, 14, 17, 22, 23 Räumliche Entfernung zwischen beteiligten Forschungseinrichtungen (21), Entwicklung gemeinsamer Begriffe und Konzepte (17)
	Nachhaltigkeit der Vernetzung (1)		Wissenschaftliche Netzwerke (3)

Tabelle 6 und Tabelle 7 zu den Kriterien der praxisorientierten und interdisziplinären Forschung zeigen, dass sich die von der Ökolandbauforschern vorgeschlagenen Kriterien weitgehend mit jenen Kriterien decken, die in der Literatur zur Evaluierung inter- und transdisziplinärer Forschung genannt wurden. Die Interviewergebnisse gehen dabei über die Kriteriensets einzelner Evaluierungskonzepte hinaus und unterstreichen damit die Zielsetzung der Interviewpartner, mit dem Kriterienset die gesamte Breite der Ökolandbauforschung abzudecken. Entsprechend sind die Kriterien für inter- und transdisziplinäre Arbeitsweise in der Literatur noch wesentlich detaillierter aufgeführt, als in den Tabellen (vgl. Begusch-Pfefferkorn 2006, Bergmann et al. 2005, Stokols et al. 2003). Diese detaillierten Konzepte sind somit eine wichtige Quelle für Reflexionsprozesse innerhalb von Forschungsprojekten, die dann nachfolgend in stärker aggregierter Form in die Evaluierung eingebracht werden können.

Die Praxiseinbindung in einem Forschungsprojekt hat besonders für die Problemdefinition und Implementierung der Ergebnisse einen hohen Stellenwert (Glaeser 2007). Analog betrachtet ist die Praxiseinbindung besonders für die Entwicklung von Ziel- und Transformationswissen wichtig (Hennen, Petermann und Scherz 2004), wenn man eine Gliederung in drei interdependente Wissensarten Ziel-, System- und Transformationswissen (vgl. Pohl und Hirsch Hadorn 2006, 35) vornimmt.

Die Anwendung und Wirkung von Lösungsbeiträgen der Forschung in Praxis und Gesellschaft werden in den Interviews und in der Literatur als wichtige Evaluierungskriterien aufgeführt. Es muss dabei aber berücksichtigt werden, dass diese Wirkungen oft erst eine Zeitspanne nach Projektende erfasst werden können (time-gap). Dieses wird zwar in der Literatur thematisiert, (vgl. Defila und Di Giulio 1999, Aennis und Nagel 2003) die bestehenden Konzepte erfassen Anwendung und Wirkung jedoch kaum über das Projektende hinaus. Interviews und Workshopergebnisse unserer Untersuchung konstatieren jedoch eine elementare Notwendigkeit für eine solche **Nachbeobachtung** und geben erste Hinweise dazu, wie diese organisiert werden kann. (z. B. Dokumentation durch Forscher und Einbindung von Akteuren aus Praxis und Gesellschaft, siehe Kap. 4.5). Dabei müssen insbesondere passende Zeiträume gefunden und (wie auch bei den anderen Kriterien) externe Einflüsse auf die Anwen-

dung und Wirkung von Forschung in der Praxis berücksichtigt und die jeweiligen Zielgruppen der Forschung benannt werden.

Die Kriterien zur Bewertung von Nachhaltigkeit, gesellschaftlicher Relevanz, Systemorientierung, Vielfalt von Themen und Methoden, Innovation und Originalität und Langfristigkeit der Forschung sind weniger zahlreich und meist komplexer (vgl. Tabelle 8). Einzelne Kriterien sind aber auch sehr leicht anwendbar, bspw. für die Themen und Methodenvielfalt, die insbesondere für die Evaluierung von Forschungsprogrammen und Institutionen relevant sind.

Die Kriterien für Nachhaltigkeit, Gesellschaftliche Relevanz und Systemorientierung überlappen sich teilweise. Sie sind, wie auch die Originalität, mit etlichen Kriterien in den publizierten Evaluierungskonzepten aufgeführt. Kriterien zur Vielfalt von Themen und Methoden und Langfristigkeit der Forschung werden in der Literatur dagegen wenig berücksichtigt.

Tabelle 8: Kriterien für die Bewertung von Nachhaltigkeit, gesellschaftlicher Relevanz, Systemorientierung, Vielfalt von Themen und Methoden, Innovation und Originalität und Langfristigkeit der Forschung.

Bereich (Anzahl Personen mit Vorschlägen)	Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwendbarkeit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
Nachhaltigkeit (10/22)	Nachhaltigkeits- und Umweltkriterien (10) (vorhandene Nachhaltigkeitskriterien, mögliche bzw. tatsächliche Wirkungen, über Einbindung relevanter Akteure und Experten)	B, C	Nachhaltigkeitskriterien (13), Relevanz f. Nachhaltigkeit (12), langfristige Trächtigkeit der Ergebnisse (17)
			Gesellschaftlich relevante Produkte und ihr Beitrag zur nachhaltigen Entwicklung (5, 22, 23)
Gesellschaftliche Relevanz (14/22)	Medienpräsenz von Problemlagen oder Forschungsthemen (6) Thematisierung von Risiken, Ängsten und möglichen Fehlentwicklungen (3)	C, D	Auswertung der politisch-öffentlichen Diskussion bzw. Expertenmeinungen um gesellschaftliche Relevanz gegenwärtiger/zukünftiger Probleme abzuschätzen (14, 17)
	Auswirkungen auf öffentliche Güter (2) Gesellschaftliche Relevanz in langfristiger Beobachtung (1)	C, D	langfristige Gemeinwohlorientierung (8) Allg. Gesell. Relevanz (3, 16)
	Werte der Zielgruppen des Ökolandbaus (3)	C, D	Ist wissenschaftliche Argumentation und Managements des Prozesses problemorientiert genug (8)
Systemorientierung (10/22)	Einbezug aller notwendigen Zielgruppen des Systems bei der Problemdefinition (3) Einbezug aller Systembereiche in den Forschungsprozess (2)	C, D	Berücksichtigung möglichst vielfältiger gesellschaftlicher Dimensionen beim Zuschnitt der wissenschaftlichen Fragestellungen (17)
	Einbettung symptomorientierter Ansätze in systemorientierte Kontexte/Lösungskonzepte (2)	C, D	
	Bewertung anhand der möglichen bzw. tatsächlichen Auswirkungen (3)	C, D	
Vielfalt von Themen und Methoden (15/22)	Vielfältigkeit der Themen- und Methoden (5) Häufigkeit bisheriger Erforschung (2)	B	
	P: Wird Themen- und Methodenvielfalt gewährleistet? - durch Gremienzusammensetzung/Auswahlprozess bei der Beurteilung von Forschungsanträgen (8) oder - Durch die Budgetverteilung (4)	B, C	
	Höhere Wertschätzung von Übersichtsarbeiten/Syntheseforschung (2)	A, B	Syntheseforschung höher bewerten (2)
	Publikation in unterschiedlichen Zeitschriften (1) Inter- und transdisziplinäre Arbeitsweise (1) Methodische Vielfalt im Forschungsantrag und Projekt (1)	B	

Bereich (Anzahl Personen mit Vorschlägen)	Einzelkriterien (vollständig) (Anzahl Nennungen in den Interviews oder im Workshop (WS))	Anwend- bar- keit	Literaturquellen aus Tabelle 4 die ähnliche Kriterien auflisten + Kriterien aus der Literatur
Innovation und Originalität (13/22)	Anwendung neuer Methoden (5) Neue Themen und Fragestellungen (4) P/I: Förderung von Themen quer zu vorhan- denen Trends und Modeerscheinungen (1) Forschung „außerhalb des Systems“ (1) Originelle Lösungen für Probleme von Gesell- schaft und Umwelt (1)	B	Neue Konzepte, Methoden, Theorien zu transdisziplinärer Forschung (5, 21) Disziplinärer wissenschaftlicher Fort- schritt, Beiträge zur Erschließung wissen- schaftlichen Neulandes oder zu einem Paradigmenwechsel (17) Originalität und „Querdenkertum“ (17)
	Innovationstransfer in und mit der Praxis (2) Innovative, interdisziplinäre Arbeitsweise (2)	B, C	kreative und innovative Produk- te/Strategien/Initiierung von Diskursen (nicht nur analytische Leistungen der Wissenschaft) (2)
	Aufgriff von Forschung durch Wissenschaftsjour- nalismus (1)	B	
Langfristigkeit (10/22)	(P/I):Mittelverteilung auf langfristige und kurz- fristige Forschungsfragen	A, B	I: Vorlauftforschung und Kurzfristnachfrage gleichermaßen (8)
	Potential des Forschungsansatzes (Zielsetzungen, Hypothesen und Methoden, Arbeitsprozess) (4) bzw. Notwendigkeit (1) Potenzielle Praxisrelevanz der Zwischenergebnisse (1)	B, C	
	Qualitativ hochwertige der Außendarstellung während der Projektlaufzeit (2)	B	
	Geringere Anforderungen an die Quantität des Output (2) bzw. höhere Gewichtung von Publikation- en (2)	A, B	Syntheseforschung höher bewerten (2)
	Wirkung der langfristigen Forschung durch Nachbeobachtung (1)	B, C	

Die Kriterien in Tabelle 8 stellen eine wichtige Ergänzung, z. T. auch Konkretisierung zu den Kriterien für die Praxisorientierung und die Interdisziplinarität dar. Sie lassen sich gut in diese eingliedern, u.a. da sie mehrfach auf die mögliche oder tatsächliche Wirkung der Forschung in Praxis und Gesellschaft verweisen, wofür eine Nachbeobachtung und eine Einbindung relevanter Akteure aus Praxis und Gesellschaft in die Evaluierung notwendig ist.

So entsteht durch die Kriterien in den Tabellen ein rundes Bild für die Bewertung praxisorientierter Ökolandbauforschung. Bei aller Vielfalt die dieses Kriterienset bietet, kann sehr gut eingegrenzt werden, welche Leistungen der Forschung für Praxis und Gesellschaft für die Evaluierung relevant sind. Dieses sind sowohl der Forschungsprozess und die konkreten Veröffentlichungen und Produkte als auch die tatsächliche Relevanz und Wirkung der Forschung für Praxis und Gesellschaft (Nachbeobachtung) (Abbildung 4). Dabei wird auch die Verfügbarkeit der für die Bewertung notwendigen Informationen zu verschiedenen Evaluierungszeitpunkten dargestellt, denn eine qualitative, faktenbasierte und verständliche Bewertung erfordert eine solide Datenbasis. Diese umfasst auch eine Berücksichtigung der internen und externen Rahmenbedingungen und Einflüsse auf den möglichen Nutzen der Forschung für Praxis und Gesellschaft. Die Zielgruppen und Betroffenen der jeweiligen Forschungsvorhaben müssen nicht nur benannt werden, sondern die verschiedenen Sichtweisen von Praxis, Gesellschaft und Wissenschaft müssen auch in den Bewertungsprozess integriert werden.

Kontext: Ziele, Ressourcen, interne und externe Rahmenbedingungen der Forschung



Abbildung 4: Übersicht über die Evaluierungskriterien und die Verfügbarkeit der für die Bewertung benötigten Informationen in verschiedenen Phasen des „Projektzyklus“

4.5 Vorgeschlagene Evaluierungsinstrumente im Kontext der Literatur

4.5.1 Spektrum der Methoden für die Bewertungsgrundlagen und den Bewertungsprozess

Entscheidend für die Qualität einer Evaluierung sind nicht nur die inhaltliche Eignung der Kriterien, sondern auch die Bewertungsgrundlagen (Materialien und Informationsquellen) und die Gestaltung des Bewertungsprozesses (beteiligte Personen und Bewertung). Darüber hinaus gibt es sowohl in den Interviews als auch in der Literatur Vorschläge für eine partizipative Entwicklung des Kriteriensets mit den Evaluierten (Aennis und Nagel 2003, Stokols et al. 2003, Bergmann et al. 2005, Glaeser 2007, Albrecht et al. 2008). Es werden auch Erfolgskriterien, die von den Evaluierten vorgeschlagen werden, als sinnvoll erachtet (Aennis und Nagel 2003, Bergmann et al. 2005). Die Bewertung der Wichtigkeit von Kriterien durch die Evaluierten (Blackstock, Kelly und Horsey 2007) oder durch die Evaluierten und in Zusammenarbeit mit Vertretern der Praxis wird ebenfalls vorgeschlagen bzw. verwendet (Pedersen et al. 2009).

Vorschläge zu Evaluierungsinstrumenten wurden in den Interviews nur punktuell geäußert, daher wird in der Tabelle 9 nur eine grundsätzliche Kennzeichnung vorgenommen. Im Workshop wurden insbesondere zwei Instrumentarien, die strukturierte, datenbankgestützte Selbstdokumentation und die Einbindung der Praxis vertieft.

Tabelle 9: Bewertungsgrundlagen und Bewertungsprozesse für die Evaluierung – Vorschläge aus den Projektergebnissen (x) und der Literatur (Zahlen aus Tabelle 4).

Bewertungsgrundlagen: Materialien und Informationsquellen			Interviews, Workshop	Literaturquellen aus Tabelle 4
Datengewinnung innerhalb der Wissenschaft	Vorhandene Dokumente	Projektanträge	x	3
		Zwischen und Endberichte	x	4, 6, 10, 12, 15, 18, (z. T. auch weitere Dokumente)
		Auswertung/Dokumentationen von Forschungsprogrammen	x	4, 18, 20
		Öffentliche Medien	x	6
	erstellt durch Wissenschaftler für die Evaluierung	Strukturierte Dokumentation in einer Datenbank	x	
		Selbstberichte, Selbstdokumentation von Wissenschaftlern/ Forschungsgruppen	x	feststehende Fragen (17, 20, 21); Freie Formulierung von Zielen und Selbstbild (20); SWOT (10)
		Wirkungsketten und Szenarien (in Projektanträgen und Endberichten)	x	
		Verwertungspläne (=zu Resonanz, Anwendung und Wirkung in der Praxis ca. 3 Jahre nach Projektende, verpflichtend ggü. Forschungsförderung)	x	In Verbindung mit Publikationen siehe Niederkrotenthaler, Dorner und Maier (2011)
	Erstellt durch Externe oder Evaluatoren	Interviews mit evaluierten Wissenschaftlern	x	4, 10, 15, 16, 18, 20, 21, 6; internet questionnaire (21)
		Prozessbegleitung / Mentoring		4, 17,18, 10
		Workshops / Fokusgruppen mit evaluierten Wissenschaftlern		4, 18, 21
		Experteninterviews / Dephi-Befragungen	x	4, 9,16
		Selbstreflektion der Evaluatoren zu eigenen Werten, Zielen, Kriterien- u. Methodenauswahl		6
		Externe Analysen: - Szenarien zur Wirkung von Forschung in Praxis und Gesellschaft - Schwachstellenanalyse des Ökolandbaus - Checklisten für Ängste, Bedenken, Risiken von Forschungsthemen - Soziale Netzwerkanalyse/Soziometrie zu Co-Autorenschaften (f. Interdisziplinarität)	x	
	Datengewinnung in Praxis und Gesellschaft durch Externe oder Evaluatoren	Interviews	x	6, 11, 15, 16, 20
Praxis-Beirat, z. B. als Projekt-, / Programm-, oder instituts-eigene Struktur		x	4	
Fokusgruppen Workshops/Gruppendiskussionen (z. B. mit Praxis, Beratung, Konsumenten)		x	4; 15, Einbezug regionaler Akteure (1)	
Umfeldanalyse zu relevanten Akteuren und der Rahmenbedingungen für die Praxis		x	20	
Internetbefragung/ Panel		x		
Review für Praxis-Veröffentlichungen		x		
Ranking/Qualitätsbeurteilung von Praxis-Fachzeitschriften (4 bedingt befürwortet)		x		
Auswertung praxisorientierter Tagungen, Websites, Zeitschriften		x		
Dialogprozesse zwischen Wissenschaft und Praxis (dauerhaft, verbindlich, repräsentativ, fachübergreifend)	x			

Bewertungsprozess		Interviews Workshop	Literaturquellen aus Tabelle 4
Extern	Gutachter/ Peer-Review (nur Wissenschaft, aber teilw. mit inter- bzw. transdisziplinärer Forschung vertraut)	x	3, 4, 9, 10, 14,19, 20
	Gutachter (Wissenschaftler sind vorher Mentoren im Forschungsprozess)		17
	kooperative Gutachten (Wissenschaftler erstellen gemeinsam einen Sammelband)		13
	Gutachter, Wissenschaftler aus anderen Fachrichtungen (Blick von außen)	x	14
	Gutachter/ Peer-Review (Wissenschaft + Praxis)	x	12
	Gremien aus Wissenschaft und Praxis	x	4, 9,11,16,
	Gremien (Interdisziplinär) o. (repräsentativ für Nachhaltigkeit oder Gesellschaft) (teilw. Wissenschaft + Praxis)	x	
Intern	Diskursive Evaluation z. B. Evaluierungstagung für Programme, Institutionen und interdisziplinäre Großprojekte (Wissenschaft + Praxis)	x x	2, 5, 9, 13, 20, 22, 23 (generelle Diskussion 15)
	Diskursive Evaluation	x	2, 5, 9, 13, 22, 23
	Selbstevaluation	x	5, 9, 10,11,20,21
	Laufendes Controlling durch Programmleitung (und Advisory Board)		4, 19

Interviewergebnisse und Literatur zeigen grundsätzlich eine gute Übereinstimmung in den Evaluierungsinstrumenten. In der Literatur werden überwiegend Multimethodenansätze angewendet oder vorgeschlagen, d. h. verschiedene Beurteilungsgrundlagen und Bewertungsprozesse werden kombiniert. Bei der Datengewinnung liegt der Schwerpunkt auf Interviews mit Wissenschaftlern und der Auswertung von Endberichten. Einige Autoren nennen als mögliche Vorgehensweisen darüber hinaus die Selbstdokumentationen, die Begleitung des Forschungsprozesses durch Mentoren oder Workshops bzw. Fokusgruppen. Ein Einbezug von Vertretern der Anwendungspraxis und der Gesellschaft wird, im Gegensatz zu den Interviewergebnissen, dagegen nur selten diskutiert.

Bei den publizierten Bewertungsverfahren liegen die Schwerpunkte auf der Begutachtung durch Wissenschaftler, der Selbstevaluation und der diskursiven Evaluation. Die Kombination von interner und externer Perspektive wird in einigen Konzepten als wichtige Voraussetzung für eine adäquate Bewertung hervorgehoben (Schiller et al. 2005, Blackstock, Kelly und Horsey 2007, Walter et al. 2007, Defilla, Di Giulio und Scheuermann 2008). Die diskursive Evaluation wird auch in den Interviews stark befürwortet (Q 15/16). Diskursive Evaluation meint dabei eine Reflexion der Evaluierungsergebnisse in einem konstruktiven Diskurs von Evaluierten und Evaluatoren (vgl. Bergmann et al. 2005).

Zwei Ergebnisse aus Interviews und Workshop setzen sich deutlich von der Literatur ab: Die Datenerhebung mittels strukturierter, datenbankgestützter Dokumentation eigener Leistungen durch die Forscher wurde als besonders geeignet erachtet (9/10 WS). Gleichzeitig wurde auch die Einbindung von Akteuren aus der Praxis und Vertretern gesellschaftlicher Interessen stark befürwortet (Q 15/16) und dafür eine Reihe von Instrumenten vorgeschlagen, wie in Tabelle 9 deutlich wird.

Die Dokumentation der Leistungen für Praxis und Gesellschaft bildete einen weiteren Forschungsschwerpunkt des Projektes und wird ausführlich in Kap. 5 dargestellt.

4.5.2 Vertiefung: Praxiseinbindung

Für die Evaluierung praxisorientierter Ökolandbauforschung wurde von den Interviewpartnern der Anspruch formuliert, Sichtweisen insbesondere aus der Praxis aber auch aus der Gesellschaft zu integrieren. Dafür wurden in den Interviews und im Workshop aber auch in der Literatur verschiedene Instrumente vorgeschlagen (vgl. Tabelle 9). Damit sind Instrumente für die Einbindung von Praktikern und Vertretern der Gesellschaft vorhanden und können für die Kriterien die in Tabelle 6) mit C⁸ gekennzeichnet sind eingesetzt werden. Des Weiteren sind Methoden für den Dialog zwischen Wissenschaft und Praxis, auch über die vorgeschlagenen Evaluierungsinstrumente hinaus, grundsätzlich vorhanden: auf Projektebene z.B. in der transdisziplinären Forschung (z. B. Gerber, Hoffmann und Thomas 2009, Bergmann et al. 2010) und übergeordnet sowohl in vielen Bereichen industrienaher Forschung bereits seit Jahrzehnten etabliert als auch in Form partizipativer, international eingesetzter Verfahren zur Technikfolgenabschätzung beschrieben (Hennen, Petermann und Scherz 2004).

Für die spezifische Anwendung der vorhandenen Instrumente und Kriterien in der Forschungsevaluierung sind diese die mit den jeweiligen Zielgruppen zu erproben und weiterzuentwickeln.

In den Interviews und im Workshop wurde insbesondere auf die richtige Auswahl geeigneter Personen und eine sorgfältige Prozessgestaltung eingegangen, um die Herausforderung der Einbindung nicht-wissenschaftlicher Akteure zu meistern. Aus den Ergebnissen wurden die nachfolgenden Empfehlungen dazu erarbeitet. Einen Kurzüberblick bietet Abbildung 5.



Abbildung 5: Wichtige Eckpunkte für eine gelingende Einbindung von Akteuren aus Praxis und Gesellschaft in die Forschungsevaluierung.

4.5.2.1 Anforderungen an die Zusammensetzung der Gruppe

a) Die Gruppe darf nicht zu groß sein (25 Personen wurden als sehr groß bezeichnet) (1 WS)

⁸Kriterien, für die eine Einbindung von nicht-wissenschaftlichen Akteuren erforderlich bzw. sinnvoll ist.

- b) Angemessenes Verhältnis bzw. eine Entscheidung zwischen Dauerhaftigkeit der Gruppe und Unabhängigkeit der beteiligten Akteure
- Dauerhafte Gruppen ermöglichen eine Annäherung an gemeinsame Begriffe und Grundverständnis (1 WS), welches vor der Evaluierung erlernt werden muss (1). Die Aktivierung immer neuer Akteure ist aufwändig (1 WS) und fehlende Kenntnisse konstanter Gruppen können auch über Einzelinterviews/Konsultationen mit Experten ausgeglichen werden (1 WS).
 - Unabhängigkeit der Bewertung ist gefährdet, wenn immer wieder die gleichen Akteure beteiligt sind (1 +2 WS) Stichwort: „*Verbeamtete Praktiker*“ (B18:56). Die Anzahl an Praxisakteuren ist mittlerweile so hoch, dass es gelingen wird, auch geeignete „neue“ Akteure für die Evaluierung finden (1 WS)
- c) Definition und Beteiligung aller relevanter Zielgruppen (1 WS)
- Je systembezogener der Forschungsansatz ist bzw. je breiter die Zielgruppe der Forschung, desto mehr Akteursgruppen müssen auch in die Evaluierung einbezogen werden
 - bei der Interpretation der Evaluierungsergebnisse wird darauf Bezug genommen, welche Akteure an der Evaluierung beteiligt waren
- d) Die Vielfältigkeit/Unterschiedlichkeit innerhalb der einzelnen Akteursgruppen wird berücksichtigt (Regionen, Standorte, Betriebsgrößen, Einstellungen/Nähe zu den Prinzipien des Ökolandbaus, Ausbildungsniveau, Verbandsstruktur, starke und geringe Bezüge zum Evaluierungsgegenstand (9 + 1 WS) z.B. indem:
- ein angemessener Querschnitt aus den entsprechenden Gruppen zusammengestellt wird
 - Teilgruppen, die aus Sicht des Forschungsprojektes oder –programms als besonders relevant sind, gezielt ausgewählt werden (und diese Auswahl transparent dargestellt wird)
- e) Vielfältige Zusammensetzung der Gruppe hinsichtlich:
- gemischte Altersstrukturen in der Gesamtgruppe (1 WS)
 - ausgeglichenes (ggf. dem Sachverhalt angemessenes) Geschlechterverhältnis (Gender sensible Bewertungen vgl. Bergmann 2005)
- f) Vertretung wichtiger Kompetenzen:
- Personen (Wissenschaftler/Evaluatoren), welche einen Überblick über die Literatur und Forschungslandschaft im Themenbereich (In- und Ausland) haben, u.a. um Forschungsdefizite von fehlendem Wissenstransfer zu unterscheiden (3 + 2WS)
 - Personen außerhalb der Landwirtschaft sind zusätzlich sinnvoll für Programmevaluationen um Neuerungen und Innovationen zu entdecken und Zukunftswege aufzuzeigen (2 + 1 WS)

4.5.2.2 Anforderungen an die Akteure aus Praxis und Gesellschaft

(Der Fokus der Interviewpartner und Workshopteilnehmer richtete sich hier insbesondere auf die landwirtschaftliche Praxis, Berater und Vertreter von Verbänden und Organisationen.)

- a) Interesse am Thema bzw. Evaluierungsgegenstand (2)
- b) Entsprechende (Detail-)kenntnisse über Thema bzw. Evaluierungsgegenstand (z.B. durch Beratung oder Betriebsschwerpunkt im Forschungsbereich) (4)
- c) Die Akteure sind möglichst intensiv im Projekt/Evaluierungsgegenstand involviert und/oder hatten ausreichend Möglichkeiten, sich mit den Forschungsergebnissen auseinanderzusetzen (1). Die jeweilige Tiefe der Auseinandersetzung muss bei der Interpretation der Evaluierungsergebnisse berücksichtigt werden.
- d) **Gleichzeitig** aber auch eine Unabhängigkeit gegenüber den Wissenschaftlern, Projekten oder Institutionen (1 WS)
- e) Betroffenheit gegenüber dem Thema (z.B. potenzielle Anwender oder von den Auswirkungen betroffen) (B4:12, B23:15) **und gleichzeitig:**
- f) Genügend Distanz zur persönlichen Situation/betrieblichen Problemlagen (5)
- g) Weitblick für zukünftige Entwicklungen, Einflüsse und Rahmenbedingungen im zu evaluierenden Themenbereich (6)
- h) Offenheit für Neuerungen (3)
- i) Sie sind in ihren kommunikativen Kompetenzen für eine Bewertung von Forschungsergebnissen geeignet (verfügen ggf. schon über Erfahrungen im Austausch zwischen Wissenschaft und Praxis) (4 + 1 WS)
- j) Die Weiterentwicklung des Ökolandbaus, ist für die beteiligten Akteure ein wichtiges Anliegen. Hierzu besteht eine Übereinstimmung zwischen den Zielen des zu evaluierenden Forschungsprogramms und den Zielen der Evaluatoren und beteiligten Akteure (3). Dies schließt eine zusätzliche Einbindung von Kritikern jedoch nicht aus.

4.5.2.3 Anforderungen an die partizipativen Prozesse zur Evaluierung angewandter Forschung

- a) Der Prozess zur Auswahl der beteiligten Praxisakteure orientiert sich an Auswahlkriterien (s.o.) und wird transparent dargestellt (3).
- b) Beiträge von Praxisakteuren werden in sinnvollem Maß eingebunden und wertgeschätzt (keine Alibi-Partizipation, kein „Belächeln“ und „Bloßstellen“ der Praxisakteure) (2)
- c) Die „eigene“ Sprache, Kommunikationsweise, der Kontextbezug bzw. die „Lebenswelt“ der Praxisakteure wird im Evaluierungsprozess z.B. durch eine Moderation/Mediation berücksichtigt (4 + 1 WS).
- d) Gezielte Entscheidung für Gruppen oder Einzelakteure:

- Um Einzelpersonen nicht zu überfordern oder zu überfrachten: Je weniger die beteiligten Praxisakteure mit der Wissenschaft vertraut sind desto eher müssen statt Einzelpersonen Arbeitsgruppen von Praxisakteuren eingebunden werden (Empfehlung Autorenteam).
 - Die Einbindung von Gruppen mindert den Einfluss der spezifischen Problemlagen und Sichtweisen einzelner Betriebe (1 WS).
 - Akteure, welche in der Wertschöpfungskette eher als Konkurrenten auftreten, sind über Interviews und nicht über Arbeitsgruppen einzubeziehen (1 WS).
- e) Zielgruppenspezifische Anforderungen an den Aufwand: der Aufwand für die Vorbereitung muss bei landwirtschaftlichen Praktikern kleiner sein (z.B. nicht 50 sondern nur 10 Anträge á 7-10 Seiten) als bei Beratern und Mitarbeitern von Verbänden (1 WS)
- f) Zeitpunkt der Sitzungen oder der Arbeitsphasen richtet sich nach den Bedürfnissen der beteiligten Praxisakteure (z. B. nicht während landwirtschaftlicher Arbeitsspitzen) (1 WS)
- g) Praxisakteure werden auf ihre Aufgabe vorbereitet und gezielt dafür qualifiziert (1 WS)
- h) Die an der Evaluierung beteiligten Praxisakteure bewerten die Forschung mit Hilfe transparenter Kriterien (um bspw. Popularität von prinzipieller Anwendbarkeit zu unterscheiden) (2).
- i) Honorierungssysteme für den Evaluierungsaufwand der Praxisakteure, welche eine zeitlich angemessene Auseinandersetzung mit den Forschungsergebnissen bzw. dem Evaluierungsgegenstand ermöglichen und rechtfertigen (2).
- j) Die Gestaltung der Evaluierungsabläufe ermöglicht intrinsische Motivation, d. h. die beteiligten Akteure können einen persönlichen Nutzen aus dem Evaluierungsprozess ziehen (z.B. Lerneffekte) (1 WS).
- k) Honorierung der Arbeit der Evaluatoren (Wissenschaftler) (z.B. Ausgleich des Zeitaufwandes, adäquate Anerkennung als berufliche Leistung, finanzielle Honorierung) (1 WS).

4.6 Zwischenfazit aus Projektphase I und II - Identifikation von Möglichkeiten der breiteren Etablierung

Die Literaturstudie hat gezeigt, dass für die Bewertung angewandter, inter- und transdisziplinärer Forschung bereits eine Fülle von Evaluierungskriterien existiert. Im Rahmen von Interviews zur angemessenen Evaluierung der Ökolandbauforschung wurden vergleichbare Kriterien gefunden und Kriterien anderer Autoren um spezifische, neue Kriterien ergänzt.

Trotz umfassender und teilweise auch erprobter Konzepte für die Evaluierung praxis- und nachhaltigkeitsorientierter, inter- und transdisziplinärer Forschung findet keine breitere Anwendung im Sinne einer grundsätzlichen Ergänzung der etablierten Forschungsevaluierung statt. Obwohl vor allem die quantitativen Kriterien grundsätzlich relativ leicht anwendbar sind. Daher stellt sich die Frage, was zu tun ist, um die derzeit übliche Forschungsevaluierung verstärkt durch praxis- und nachhaltigkeitsorientierte Kriterien zu ergänzen.

Aus Sicht der Autoren sind dafür die folgenden Punkte relevant:

- **Laufende Aktivitäten** sollten gebündelt werden, um **vorhandene Konzepte** in der Scientific Community zu diskutieren. In den Konzepten besteht große Übereinstimmung bezüglich der Leistungen, die zu berücksichtigen sind. Deshalb sollte die Auswahl von Kriterien für einen möglichst fachübergreifenden „Minimalstandard“ zur Evaluierung von Forschungsleistungen für Praxis und Gesellschaft machbar sein. In die Entwicklung eines erweiterten Evaluierungsverfahrens, das auch Leistungen für Praxis und Gesellschaft bewertet, sollten auch Entscheidungsträger aus Forschungspolitik und Forschungsförderung eingebunden werden, damit gleichzeitig der Weg für dessen breite Einführung bereitet wird.
- **Verbesserte Umsetzbarkeit einer erweiterten Evaluierung:** Die bestehenden Einzelkonzepte arbeiten derzeit mit relativ aufwändigen Methoden der Datenerhebung (z. B. Interviews, Dokumentenanalyse etc.) und die gewonnenen Daten werden nicht für verschiedene Evaluierungsgegenstände und –zeitpunkte verfügbar gemacht. Eine Verbesserung der Datengrundlage zu Leistungen der Forschung für Praxis und Gesellschaft mit dem Ziel, deren Evaluierung zu erleichtern, wird dabei bisher erst von einigen wenigen Autoren geäußert, z. B. Uriate et al. (2007) und SIAMPI (2010). Hier wird aus Sicht des Projektes jedoch ein entscheidendes Entwicklungspotenzial gesehen: Durch die Integration der Datenerfassung für die Evaluierung in bereits bestehende Dokumentationsverpflichtungen könnte der zusätzliche Aufwand für die Datenerfassung minimiert und die Datenverfügbarkeit deutlich verbessert werden. Dies ist eine wichtige Grundlage für eine breite Akzeptanz einer erweiterten Evaluierung unter Wissenschaftlern, wie die Diskussionsergebnisse innerhalb der Transferphase des Projektes zeigten.

5 Konzeptentwicklung für die strukturierte Dokumentation von Leistungen für Praxis und Gesellschaft in Anträgen und Berichten der öffentlichen Forschungsförderung

Aufbauend auf den Erkenntnissen aus Projektphase I und II (siehe Kap. 4.6) wurde vertieft an einem Konzept für die Verbesserung der Datengrundlage für die Evaluierung gearbeitet. Diese Konzeptentwicklung umfasst:

- Eine kurze Analyse der bestehenden Forschungsdokumentation und ihrer Nutzung für Evaluation und andere Zwecke. Recherche von vorhandenen Initiativen für eine Vereinfachung und Standardisierung der Datenerfassung auf institutioneller Ebene und im Rahmen der Projektevaluierung.
- Ein Vergleich des erarbeiteten Datenbedarfs für die Evaluierung von Leistungen für Praxis und Gesellschaft mit den Dokumentationsanforderungen in Anträgen und Berichten für die Projektträger der deutschen Ministerien, unter Berücksichtigung spezifischer Zielsetzungen der Programme des BMELV.
- Die Entwicklung eines Konzeptes für eine in Anträge und Berichte integrierte Erfassung von Daten und deren anpassungsfähige, flexible Nutzung für verschiedene Evaluationen.

5.1 Die vorhandene Dokumentation von Forschungsleistungen für die Evaluierung

Derzeit werden wissenschaftliche Leistungen (Abbildung 6) in erster Linie in Veröffentlichungen, Berichten und Anträgen dokumentiert. Davon sind bisher Veröffentlichungen mit Peer-Review und darauf aufbauende Kennzahlen (z. B. SCI) für die Evaluierung des Scientific Impact besonders relevant. Die Dokumentation in Anträgen und Berichten dient dagegen fast ausschließlich der Drittmittelakquisition, deren Erfolg in Form des Drittmittelvolumens wiederum als Kriterium der etablierten Evaluierung herangezogen wird. Gleichzeitig werden für die Bewertung der wissenschaftlichen Qualifikation der Antragsteller i. d. R. die Indikatoren der etablierten Evaluierung herangezogen.

Die auf Peer Review Publikationen aufbauenden Informationen können über Thompson Reuters⁹ und andere Datenbanken abgefragt werden. Weitere Informationen die für die Evaluierung von Forschung herangezogen werden wie z. B. das Drittmittelvolumen, Publikationen ohne Peer Review und Konferenzbeiträge oder auch Leistungen in der Lehre oder Selbstverwaltung werden in der Regel auf institutioneller Ebene erfasst. Sie sind dementsprechend individuell zugeschnitten, d.h. sehr heterogen und mit unterschiedlicher Tiefenschärfe, wodurch sich für Institutionenübergreifende Evaluationen erhebliche Probleme ergeben.

⁹http://thomsonreuters.com/products_services/science/science_products/scientific_research/research_evaluation_tools/

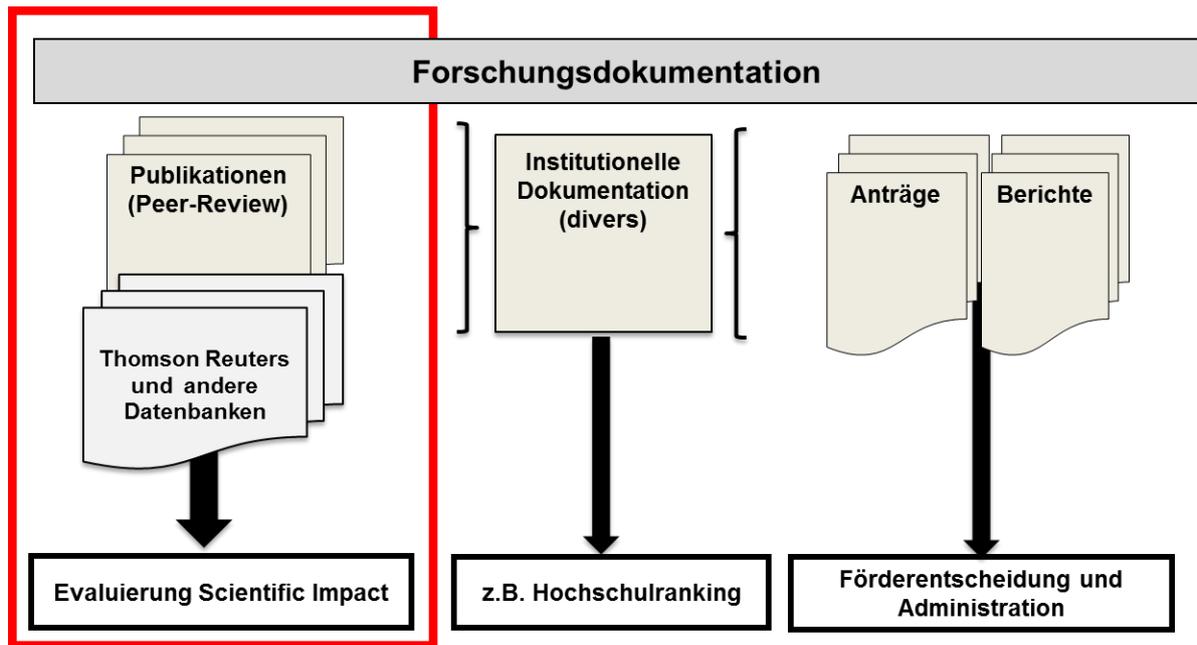


Abbildung 6: Die bisherige Forschungsdokumentation und ihre Nutzung für Forschungsevaluierung und Drittmittelakquise.

Eine Vereinfachung und Standardisierung der Datenerfassung für die Forschungsevaluierung ist daher derzeit ein wichtiges Thema, sowohl national als auch auf europäischer Ebene. Im Rahmen der Pilotstudie zu einem einheitlichen Forschungsrating ist eine Unterarbeitsgruppe des Wissenschaftsrates mit der Entwicklung von Standards der Datenerhebung beschäftigt. Ziel ist es dabei, den Aufwand der wissenschaftlichen Einrichtungen für die Dokumentation zu begrenzen, indem sich die Akteure, die für Evaluationszwecke Forschungsdaten erheben, über die Erhebungsformate stärker abstimmen und so eine Mehrfachnutzung, der Daten ermöglichen (Wissenschaftsrat 2010, 10). Der Fokus des Forschungsratings liegt dabei auf wissenschaftlichen Leistungen, jedoch wird auch die Dimension Transfer mit berücksichtigt (Wissenschaftsrat 2009).

Ähnliche Entwicklungen finden sich auf europäischer bzw. internationaler Ebene: So wird für die Erfassung der europäischen universitären Forschung für Hochschulrankings festgestellt, dass „*assessment of university-based research is being hampered by a lack of reliable, comparable, and comprehensive data.*“ Daraus folgernd wird eine eigenständige, multidimensionale, flexible Matrix für die webbasierte Datenerfassung erforderlich gehalten (AUBR EG 2009, 15–16). Für das Ranking werden nicht nur die klassischen wissenschaftlichen Performanzindikatoren vorgeschlagen, sondern auch Forschungsethik und gesellschaftlicher und ökonomischer Nutzen, wobei in diesen Bereichen eine weitere Indikatorenentwicklung für notwendig gehalten wird (AUBR EG 2009, 40–42).

Die Entwicklung eines multidimensionalen, internationalen Instrumentes für das Hochschulranking wird unter anderem in einer Machbarkeitsstudie des 7. FRP-Projektes U-Multirank entwickelt. Mit dem Ziel die Vielfalt im Hochschulbereich adäquat abzubilden bezieht sich das Ranking auf Studium & Lehre, Forschung, internationale Orientierung, Wissenstransfer und regionales Engagement. In einem Stakeholder orientierten Ansatz wurde u.a. untersucht, inwieweit europäische Universitäten Daten zu diesen Kategorien liefern können (CHERPA-Network 2010.). Festgestellt wurde, dass das Konzept für die drei erstgenannten Kategorien tragfähig ist, jedoch die vorhandenen Daten zum Wissens-

transfer oft noch nicht hinreichend sind und für „regionales Engagement“ nur einzelne Universitäten Daten bereitstellen konnten (Federkeil 2011).

Mit den Research Councils UK und dem Standard Evaluation Protocol in den Niederlanden existieren zwei Verfahren der generalisierten Datenerfassung für Leistungen der Forschung über den wissenschaftlichen Impact hinaus.

Die Research Councils UK fokussieren in ihrer Forschungsförderung auf “Excellence with Impact”. Wobei wissenschaftliche Exzellenz nach wie vor das primäre Kriterium ist, der Impact jedoch klar auf den gesellschaftlichen Nutzen referenziert. Antragsteller erstellen ein “Impact Summary” (welcher potentielle Nutzen, für wen) und beschreiben “Pathways to Impact” (Beiträge der Forschung zur Verwirklichung dieses potentiellen Nutzens), mit dem Kernziel, dass die Forscher sich mit dieser Thematik auseinandersetzen (RCUK 2010). Gleichzeitig ist den Research Councils dabei bewusst, dass sich Impact in jeder Forschungsphase und darüber hinaus zeigen kann, sowie unvorhergesehene und nicht intendierte Wirkungen zustande kommen. Daher wird ein “Outcome Collection System” verwendet, um Ergebnisse und beobachteten Impact auch noch bis drei Jahre nach Ende eines Projektes erfassen zu können (RCUK 2010, RCUK 2011a). Die Datenerfassung der Research Councils sind in ein Schnittstellenmanagement mit den Higher Education Institutions eingebunden (RCUK 2010, 2011b) und werden für das Research Excellence Framework (REF) verwendet. Das REF dient dazu, die staatlichen Mittel anhand von Leistungsindikatoren an die Universitäten zu verteilen. Dieses Assessment System wurde aus dem Research Assessment Exercise (RAE) entwickelt und hat über die Zeit an Größe und Komplexität kontinuierlich zugenommen. Aus diesem Grund wird der Nutzen gegenüber dem Aufwand zunehmend in Frage gestellt, wie auch die auf diesem System basierende Mittelverteilung an die Universitäten generell (Martin 2011).

Das Standard Evaluation Protokoll für die Evaluierung Niederländischer Universitäten umfasst u.a. die Erfassung von Publikationen für das Fachpublikum und die breitere Öffentlichkeit, sowie ein Sepsreport zur Sozialen Relevanz des Institutes.

5.2 Integration der Datenerfassung in die Antragsstellung und Berichterstattung.

Grundlage für das etablierte System zur Evaluierung des Scientific Impact ist die Dokumentation von Forschungsergebnissen in wissenschaftlichen Papers. Im Rahmen des Peer-Review sind Wissenschaftler dazu bereit die Arbeit von Kollegen mit entsprechendem zeitlichem Aufwand zu bewerten und damit zur wissenschaftlichen Qualitätssicherung beizutragen. Die Publikationstätigkeit und deren Qualitätssicherung durch das Peer-Review werden gleichzeitig für die Forschungsevaluierung verwendet, wodurch der damit verbundene Aufwand effizient genutzt wird.

Es ist deshalb naheliegend, für die Evaluierung der Leistungen von Forschung für Praxis und Gesellschaft ebenfalls etablierte Formen der Forschungsdokumentation nutzbar zu machen und dadurch über die laufenden Anforderungen hinausgehenden, zusätzlichen Dokumentationsaufwand für die Forscher zu vermeiden.

Vor dem Hintergrund, dass Daten für die Leistungen der Forschung für Praxis und Gesellschaft auf institutioneller Ebene bisher nur unzureichend erfasst werden und ohne zusätzlichen Aufwand wahrscheinlich auch nicht besser zu erfassen sind, wurde untersucht, inwieweit Anträge und Berichte für

eine Dokumentation von Leistungen der Forschung für Praxis und Gesellschaft nutzbar gemacht werden können.

Tabelle 10 zeigt links die Anforderungen der Programme des BMELV an Anträge und Berichte (BLE, BMBF 2006). Sie sind mit den Anforderungen der Ministerien BMBF, BMWi, BMFSFJ, BMELV, BMU hinsichtlich administrativer Angaben identisch. Unterschiede bestehen lediglich in den Vorgaben und Zielsetzungen der Programme (vgl. Tabelle 3) auf welche Antragsteller z. B. in den Forschungszielen oder der geplanten Verwertung Bezug nehmen müssen. Hier wird eine erhebliche Überschneidung mit dem Datenbedarf für die Evaluierung gesehen (Tabelle 10, rechts ausführlich in Anhang III). Beispielsweise wird die Dokumentation von Arbeitsschritten auch Praxiskontakte oder interdisziplinäre Arbeitsweisen einschließen, wenn sie für ein Projekt relevant waren. Die Darstellung der (geplanten) Verwertung von Projektergebnissen beinhaltet in der Regel Angaben über Veröffentlichungen und Produkte für Praxis und Gesellschaft und kann bei ausreichender Detailliertheit der Bewertung von potenzieller Praxisrelevanz dienen.

Tabelle 10: Anforderungen für die Berichterstattung und Beantragung öffentlicher Forschungsgelder des BMBF, BMWi, BMFSFJ, BMELV, BMU, BLE (BLE, BMBF 2006) (links) und für den Datenbedarf zur Evaluierung wissenschaftlicher Leistungen für Praxis und Gesellschaft (rechts).

Informationen für die Beantragung und Berichterstattung	Antrag/Planung	Admin.	Abschluss/Bericht	Datenerfassung zur Evaluierung der Leistungen für Praxis und Gesellschaft	Antrag	Bericht
Ressourcen (€, PM)	Easy-AZA	profi	profi	Ressourcen (€, PM, Projektbeteiligte)		
Kompetenzen	AZA 6			Aus- u. Weiterbildung (Fachlich, Methodisch, Praxisbezogen)		
Ziele/Zielerreichung	AZA 6 FISA-online/e-prints	Änderungsanträge/-bescheide, Zwischenberichte	Schlussbericht + Prüfung durch Sachbearbeiter + e-prints	Ziele/Zielgruppen (Nutzer, Betroffene)		
Arbeitsschritte / Projektstruktur	AZA 6			Projektstruktur, Rahmenbedingungen für Praxisbezug		
Beitrag zu förderpolitischen Zielen				Interdisziplinarität		
Verwertung(splan)				Praxis-Kontakte		
Ergebnisse				Wirkungen in Praxis/Gesellschaft (zum Projektende u. langfristig)		
Veröffentlichungen (Wissenschaft und Praxis)				Ergebnisse		
				Veröffentlichungen und Produkte für die Praxis		
				wiss. Publikationen (Scientific Impact)		

Der notwendige inhaltliche Anpassungsbedarf in der Dokumentation ist schematisch in Tabelle 11 dargestellt. Aufgrund der guten Übereinstimmung der benötigten Informationen liegt es nahe, die Dokumentationsleistungen, die für Projekte in den genannten Programmen erbracht werden, für die Evaluierung von Leistungen der Wissenschaft für Praxis und Gesellschaft nutzbar zu machen. Hierfür wären, gemäß der Analyse zum Datenbedarf für die Evaluierung, folgende Anpassungen erforderlich:

- stärkere Formalisierung und Standardisierung der Dokumentation, um eine Auswertung der Daten sowohl nach Gesichtspunkten der Forschungsförderung, als auch der Evaluierung zu erleichtern

- explizite Dokumentationsstrukturen, um sicher zu stellen, dass bisher nur implizit erfragte Leistungen, die unter Umständen nicht mit ausreichender Ausführlichkeit und Eindeutigkeit dokumentiert wurden, künftig in jedem Fall mit hinlänglicher Detailschärfe beschrieben werden
- Ergänzung von Informationen, die innerhalb der Forschungsförderung bisher nicht eingefordert werden (z. B. Informationen zu Zielgruppen und Betroffenen)
- Ergänzung der Dokumentation mit Wirkungen der Forschung in Praxis und Gesellschaft, die auch noch nach Projektabschluss identifiziert und dokumentiert werden können
- Dokumentation der Daten in Formaten, die ein Einlesen in ein Datenbanksystem ermöglichen bzw. direkte Dokumentation in einem Datenbanksystem

Tabelle 11: Inhaltlicher Ergänzungs- und Anpassungsbedarf, um die bisherige Dokumentation in Anträgen und Berichten für die Forschungsevaluierung nutzbar zu machen.

Datenerfassung Forschungsförderung	Ergänzungen zur Evaluierung der Wirkungen in Praxis/Gesellschaft	Anpassungs-, u. Standardisierungs- bedarf	Antrag	Bericht
Ressourcen (€, PM)	Projektbeteiligte	geringfügig		
Kompetenzen	Aus- u. Weiterbildung (Praxisbezug, Kommunikation)	geringfügig		
Ziele/Zielerreichung Arbeitschritte / Projektstruktur	Zielgruppen (Nutzer, Betroffene) Rahmenbedingungen für Praxisbezug	mäßig		
Beitrag zu förder- politischen Zielen Verwertung(splan)	Praxis-Kontakte vervollständigen Wirkungen in Praxis/Gesellschaft (auch nach Projektende z.B. bei Antrag Folgeprojekt)	mäßig		
Ergebnisse	ggf. Feedback von Akteuren aus Praxis und Gesellschaft	geringfügig		
Veröffentlichungen (Wissenschaft und Praxis)	Veröffentlichungen und Produkte nach Projektende	geringfügig		

Für eine stärkere Formalisierung und Strukturierung von Anträgen und Berichten sprechen auch die Erfahrungen von Programmverantwortlichen Mitarbeitern in forschungsfördernden Einrichtungen. Sie betonten in Workshops, in denen vorläufige Projektergebnisse diskutiert wurden, dass es zur Zeit nicht oder nur mit Einschränkungen möglich ist, die Erreichung der angeführten Programmziele durch die geförderten Projekte nachzuweisen, so dass die bisherige Dokumentationsform auch schon für interne Evaluierungszwecke nicht geeignet erscheint. Auch für die Evaluierung von Programmen sind daher derzeit jeweils umfassende Aktivitäten zur Datenerhebung erforderlich. Eine Formalisierung und teilweise Standardisierung der Dokumentation von Leistungen für Praxis und Gesellschaft

im Zuge der Berichterstattung würde diesbezüglich neue Möglichkeiten eröffnen und könnte gleichzeitig ohne besonderen zusätzlichen Aufwand die Grundlage für die Evaluierung dieser Leistungen in anderen Zusammenhängen bilden (siehe Abbildung 7).

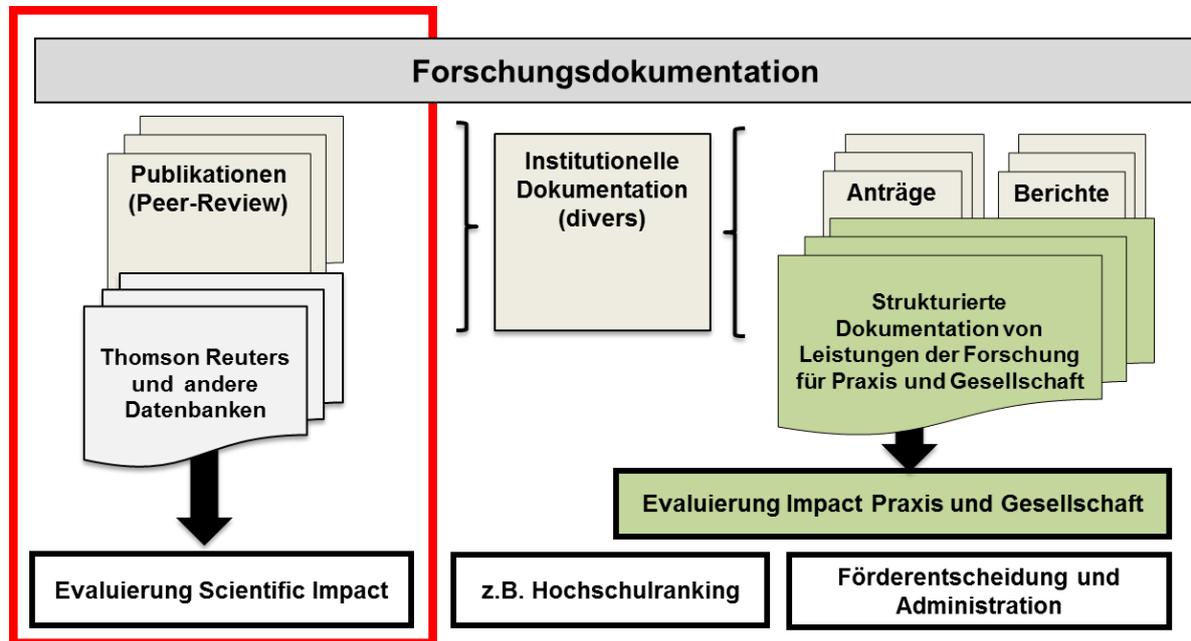


Abbildung 7: Leistungen für Praxis und Gesellschaft werden bereits jetzt in Anträgen und Berichten dokumentiert und könnten auch für die Evaluierung genutzt werden.

5.3 Konzept für die strukturierte, datenbankgestützte Dokumentation in Anträgen und Berichten

Wie die vorgeschlagene Dokumentation von Leistungen für Praxis und Gesellschaft im Detail gestaltet werden könnte, wird nachfolgend vorgestellt. Wie das Vorgehen zur Evaluierung gestaltet werden könnte und ob ein dem Peer-Review wissenschaftlicher Papers vergleichbares Verfahren möglich wäre, wird ebenfalls diskutiert.

Die Dateneingabe und Datennutzung ist in Abbildung 8 dargestellt. Hauptgrund für die **Dateneingabe** ist die **Antragstellung und Berichterstattung für Drittmittelprojekte**, jedoch kann auch eine **freie Dateneingabe** erfolgen. Darüber hinaus ist ein Schnittstellenmanagement mit Institutionen sinnvoll um den Dokumentationsaufwand für die Wissenschaftler zu reduzieren.

Die Speicherung und Verwaltung in einem **Datenbanksystem**, welches vorhandene Datenbanksysteme der Forschungsförderung ergänzt ermöglicht eine Nutzbarkeit der Daten für die Evaluierung und die Projektträger. Der gesamte **Datensatz** eines solchen Dokumentationssystems könnte (nach Freigabe durch die Dateneigentümer) **für die Evaluierung von Projekten, Programmen Institutionen und einzelnen Wissenschaftlern** verwendet werden, da die dokumentierten Leistungen nicht nur auf die Projekte bezogen sind, innerhalb derer sie dokumentiert wurden, sondern sich den verschiedenen Evaluierungsgegenständen zuordnen lassen. Für einen spezifischen Evaluierungsgegenstand können die Daten kontextbezogen gefiltert werden, je nachdem welche Kriterien als wichtig erachtet werden. Die Daten können wahlweise detailliert oder bei entsprechend gestalteter Datenabfrage stärker aggregiert dargestellt werden, um sie an die Tiefe und Umfang einer Evaluierung anzupassen. Es

steht eine qualitativ hochwertige, weitgehend vollständige und leicht auszuwertende Dokumentation sowohl für eine **interne** als auch für eine **externe Evaluierung** zur Verfügung (Abbildung 8).

Für die Akzeptanz dieses Systems bei Forschern dürften seine Anwenderfreundlichkeit und eine verminderter oder gleichwertiger Zeitaufwand für die Beantragung von Projekten und der Berichterstattung über Forschungsleistungen entscheidend sein. Wenn es darüber hinaus auch Vereinfachungen für Mitarbeiter von Forschungsförderungseinrichtungen bei der Verwaltung von Projekten in Aussicht stellt, würde dies die Bereitschaft für seine Nutzung weiter erhöhen.

Es wird hier davon ausgegangen, dass Evaluierungen anlassbezogen stattfinden. Die Auswahl der Kriterien und die Gestaltung des Bewertungsprozesses können sehr unterschiedlich ausfallen, weil neben den Auftraggebern und Evaluatoren ggf. auch die Evaluierten das Vorgehen mitgestalten. Die Datengrundlage ist für diese unterschiedliche Nutzung breit genug anzulegen und die Ausgestaltung der Evaluierung kann durch Empfehlungen zu Kriterien und Bewertungsprozessen unterstützt werden. Beispielsweise können durch eine verbesserte und leicht handhabbare Datengrundlage die Wirkungen in Praxis und Gesellschaft besser nachvollzogen und damit die Zielerreichung von Forschungsprogrammen nachgewiesen und die künftige Ausrichtung der Forschung gestaltet werden. Interne Evaluierungen können gezielt für Lernprozesse innerhalb der Organisation genutzt werden.

Eine Datenbasis, wie sie hier vorgeschlagen wird, eröffnet jedoch auch die Möglichkeit eine von einzelnen Anlässen unabhängige grundsätzliche Evaluierung (wie beim Scientific Impact) zu entwickeln und zu erproben. In Ergänzung zur Begutachtung von Projektanträgen könnte beispielsweise auch ein „Peer-Review“, der für diesen Zweck angepassten Berichterstattung über praxisorientierte Forschungsprojekte unter Einbeziehung von Praktikern und Vertretern der jeweils relevanten Bereiche der Gesellschaft erfolgen. Dabei sind jedoch auch noch viele offene Fragen zum Vorgehen oder zur Gutachterausswahl zu beantworten. Unter anderem müssen Antworten darauf gefunden werden, wie Evaluatoren aus Praxis und Gesellschaft ein unmittelbarer Nutzen vermittelt werden kann, wenn sie Forschungsleistungen ehrenamtlich bewerten sollen und/oder wie ihre Leistungen konkret honoriert werden können. Hierfür ist ein breiter Diskurs in der Scientific Community und darüber hinaus notwendig.

Ein entscheidender Vorteil des hier vorgestellten Konzeptes ist, dass die Dokumentation wissenschaftlicher Leistungen nicht ausschließlich zum Zweck der Evaluierung durchgeführt. Der Nutzen dieser Art von Forschungsdokumentation geht **für die Projektträger** über die Evaluation hinaus, weil er auch die **Administration** von Projekten erleichtern würde. Denn eine stärker strukturierte Dokumentation ermöglicht die gezielte Auswahl benötigter Informationen z. B. für einen Vergleich von Arbeitsschritten im Antrag mit Zwischen- oder Schlussberichten (Abbildung 9) und vereinfacht damit die Überprüfung der vertragsgemäßen Projektbearbeitung.

Für die Wissenschaftler könnte, bei einer angepassten Umsetzung der Dokumentation, der **Aufwand für Anträge und Schlussberichte bei entsprechender Gestaltung verringert** werden. Des Weiteren lassen sich die Informationen auch für den Wissenschaftler selbst nutzen, z. B. für **Bewerbungen oder Berufungsverfahren**.

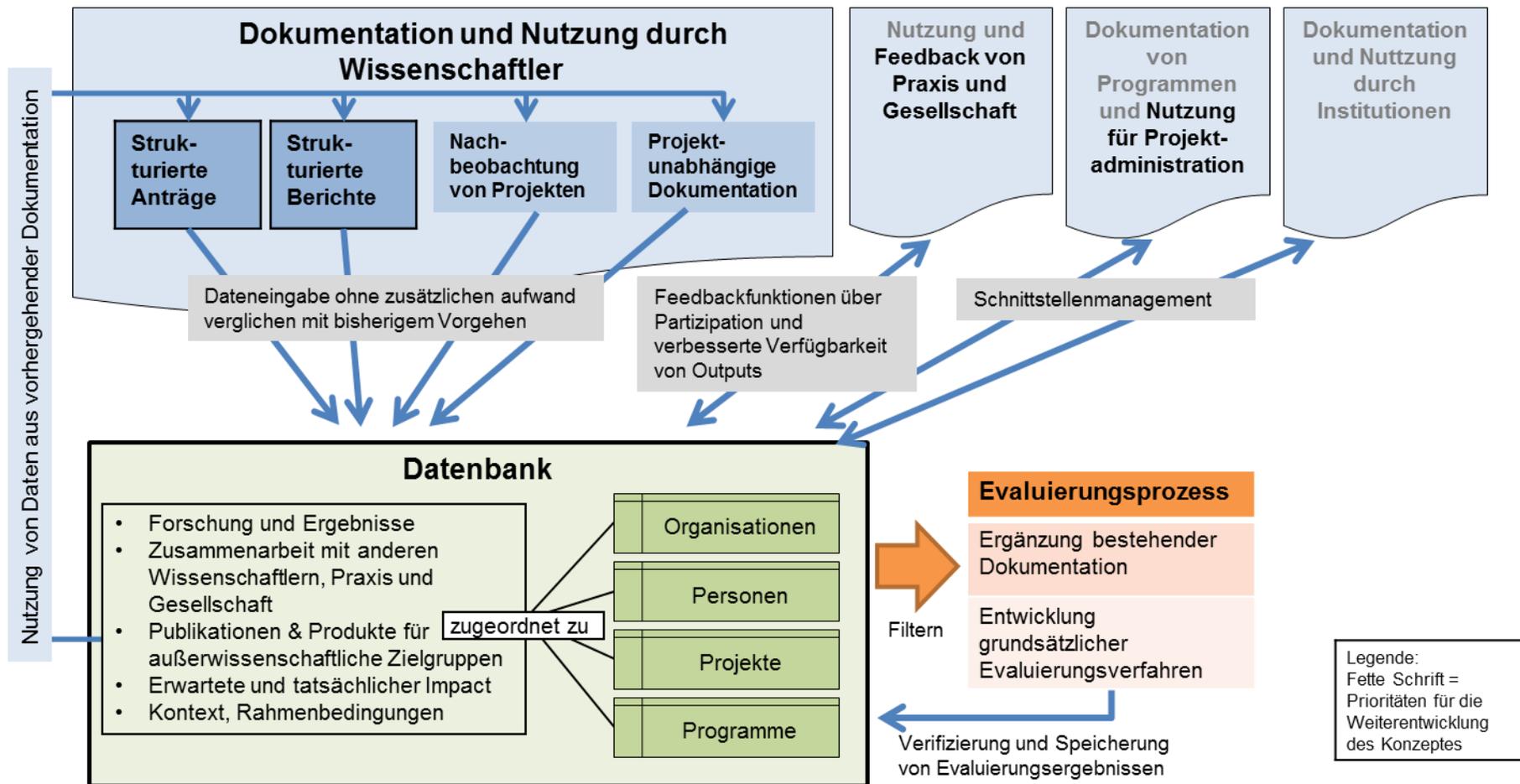


Abbildung 8: Nutzung der strukturierten Dokumentation für die Evaluierung verschiedener Evaluierungsgegenstände.

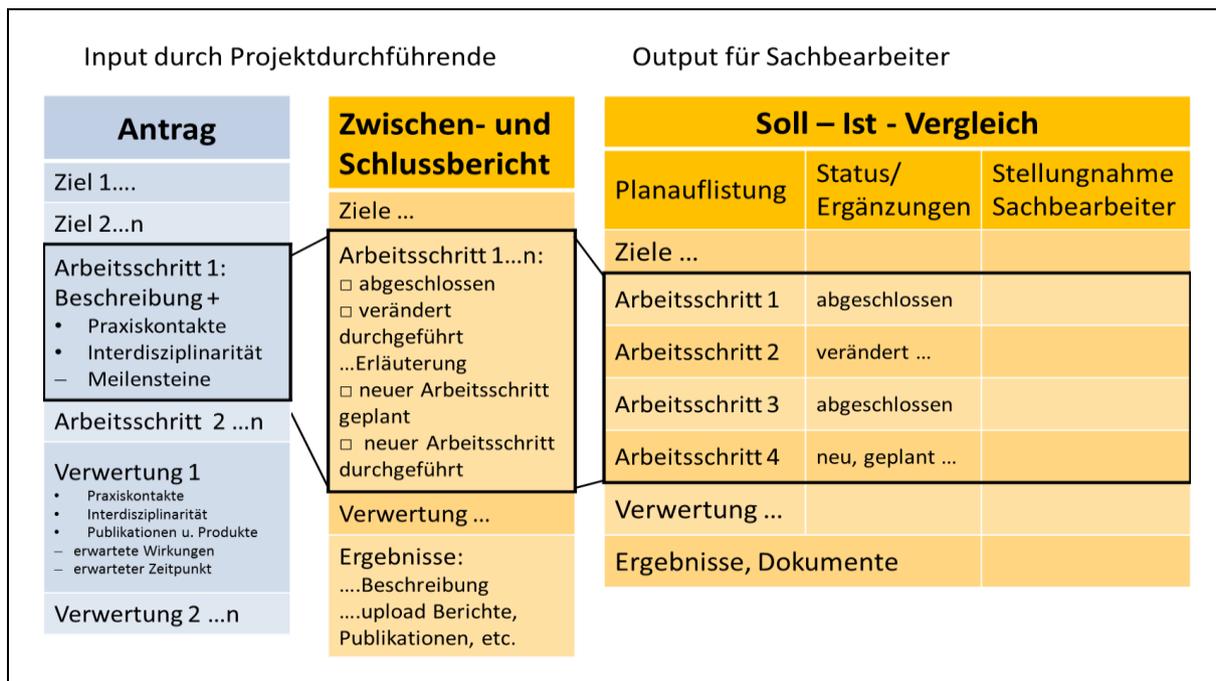


Abbildung 9: Beispiel: Vereinfachte Berichterstattung und Projekt-Administration durch strukturierte Dokumentation.

Abschließend werden in Tabelle 12 die erforderlichen Eigenschaften und die damit beabsichtigten Ziele zusammengefasst.

Tabelle 12: Übersicht über die notwendigen Eigenschaften des Dokumentationssystems und der damit beabsichtigten Ziele und Funktionen für die Evaluierung von Leistungen der Forschung für Praxis und Gesellschaft

Erforderliche Eigenschaften	Ziele und Nutzen
Systematische, strukturierte, datenbankgestützte Dokumentation	<ul style="list-style-type: none"> • Verwendbar (filterbar) für verschiedene Evaluierungsgegenstände (Wissenschaftler, Projekte, Programme, Institutionen), sowie für unterschiedliche Kontexte dieser Evaluierungsgegenstände und diverse Zielsetzungen der Evaluation
kategorisierte Informationen kombiniert mit Text/Dokumenten	<ul style="list-style-type: none"> • Transparenz und kontextbezogene Verwendbarkeit • Das System wird der Komplexität von Forschung gerecht
Offen für alle praxisrelevanten Leistungen der Agrarwissenschaft, unabhängig von der Finanzierungsquelle	<ul style="list-style-type: none"> • Schafft die Voraussetzungen für die Beteiligung der gesamten Agrarforschung • Gewährleistet eine vollständige Dokumentation für die Evaluierung von Wissenschaftlern, Institutionen, Projekten und Programmen
Eignung des Dokumentationskonzeptes für die Antragsstellung und Berichterstattung	<ul style="list-style-type: none"> • Minimierter Aufwand für Forscher und Projektadministration, • Verbesserte Entscheidungsgrundlage für die Projektadministration • Verbesserte Steuerung und Weiterentwicklung von Programmen sowie Erfolgskontrolle öffentlich geförderter Forschung • Verwendung der Dokumentation in der Forschungsförderung erbringt eine kritische Masse an teilnehmenden Forschern • Aktualisierung des Impacts kann gewährleistet werden, wenn sie z.B. als Referenz bei der Einreichung neuer Projekte eingefordert werden
Zeitlich offene Dokumentation	<ul style="list-style-type: none"> • Impacts können zu dem Zeitpunkt dokumentiert werden, zu dem sie auftreten und werden damit sichtbar gemacht • Aktualisierung des Impacts ist gewährleistet
Feedback von Akteuren aus Praxis und Gesellschaft integriert	<ul style="list-style-type: none"> • Erfassung des tatsächlichen Impacts durch Einbindung von Zielgruppen und Betroffenen • Evaluierung geht über die wissenschaftliche Sichtweise hinaus
Erprobte Nutzerfreundlichkeit	<ul style="list-style-type: none"> • Erleichterung der Dokumentation wissenschaftlicher Arbeit für Forscher • Gute Nutzbarkeit für Evaluation und Projektadministration. • Pragmatische Dokumentation, die sich auf dokumentierbare und belegbare Aktivitäten und Wirkungen der Wissenschaft konzentriert

5.3.1 Das vorgestellte Konzept im Vergleich zur etablierten Forschungsevaluierung

Ein Vergleich zwischen der etablierten Forschungsevaluierung und dem hier skizzierten Konzept zeigt, dass bestimmte Vorteile des etablierten Systems auch mit dem vorgestellten Konzept erreicht werden können (s. Tabelle 13). Dazu gehört die Nutzung vorhandener Dokumentationstätigkeiten (beim „Scientific Impact“ die Veröffentlichung von Ergebnissen mit Peer-Review) und die zentrale

Verfügbarkeit und vielfältige Nutzbarkeit der Informationen durch eine datenbankgestützte Erfassung.

Da Veröffentlichungen genauso wie Anträge und Berichterstattung ein integraler Bestandteil der wissenschaftlichen Tätigkeit sind, ist eine hohe Anschlussfähigkeit gegeben. Diese besteht darin, dass sowohl eine kritische Masse an beteiligten Wissenschaftlern erreicht werden kann als auch ein über die Evaluierung hinausgehender Nutzen zu erwarten ist.

Tabelle 13: Vergleich der Evaluierung des „Scientific Impact“ und dem vorgestellten Konzept für die Evaluierung der Wirkungen in Praxis und Gesellschaft

	Wirkungen in der Wissenschaft (statusquo)	Wirkungen in Praxis & Gesellschaft (Konzept)
Etablierung	seit 20-30 Jahren	noch erforderlich
Daten-grundlage	Veröffentlichung von Ergebnissen (Peer-Review-System)	vorwiegend Beantragung / Dokumentation von Projekten für die Forschungsförderung
Aufwand Datenerfassung	hoher Aufwand für das Peer-Review-Verfahren Kein zusätzlicher Aufwand ab der Veröffentlichung , (da Zitationen digital erfasst werden)	Kein/kaum zusätzlicher Aufwand für Dateneingabe: erfolgt im Zuge der Antragstellung/Berichtsabgabe Auswertung der Daten ohne „automatische“ Indices
Verfahren	Evaluierung nach Veröffentlichung zentral und extern (Thompson Scientific) + Dokumentationssysteme wissenschaftlicher Einrichtungen Nutzung dieser Ergebnisse für verschiedene Gegenstandsbezogene Evaluierungen	Gegenstandsbezogene Evaluierung unter Einbindung von Forschern und Praktikern auf Basis der Datenbanken von Forschungsförderung und ggf. wissenschaftlicher Einrichtungen „Selbstlaufende“ Evaluierung könnte ggf. entwickelt werden
Nutzbarkeit	leicht verfügbare Daten, Indikatoren und Indices für verschiedene Evaluierungsanlässe	leicht verfügbare Daten und Indikatoren für verschiedene Evaluierungsanlässe ; Anfrage an Forscher/Einrichtungen zur Aktualisierung und Datenfreigabe bei anstehender Evaluierung erforderlich
Spezifika	eindeutige Systemgrenzen (Publikationen); lineare, „kausale“ Wirkungsprozesse → Qualitative Bewertungen münden in quantitativen Indikatoren	weitreichende Systemgrenzen; vernetzte Wirkungsprozesse → Profil aus quantitativen und qualitativen Indikatoren mit Berücksichtigung von Kontext und Rahmenbedingungen

Ein entscheidender Unterschied bleibt darin bestehen, dass die etablierte Evaluierung in einem System (Publikationen) mit klar abgrenzbaren Prozessen (erst Peer-Review, dann Veröffentlichung, dann Zitation) stattfindet, welches letztlich auch die (nicht unumstrittene) Bildung von Indices erlaubt. Im Gegensatz dazu sind die Systemgrenzen für die Evaluierung der Leistungen für Praxis und Gesellschaft deutlich weiter und Wirkungsprozesse vernetzt (Abbildung 10). Es wird, nach dem derzeitigen

Kenntnisstand, daher davon ausgegangen, dass diese Leistungen gut in einem Profil von quantitativen und qualitativen Indikatoren darstellen lassen. Die Bildung von Indices wurde in Interview und Workshops kritisch gesehen und kontrovers diskutiert. Die Konzepte in der Literatur sehen bisher von der Bildung von Indices ab. Lediglich das Konzept von Niederkrotenthaler, Dorner, und Maier (2011) umfasst die numerische Bewertung einzelner Leistungen, die zu einem Index aufsummiert werden.

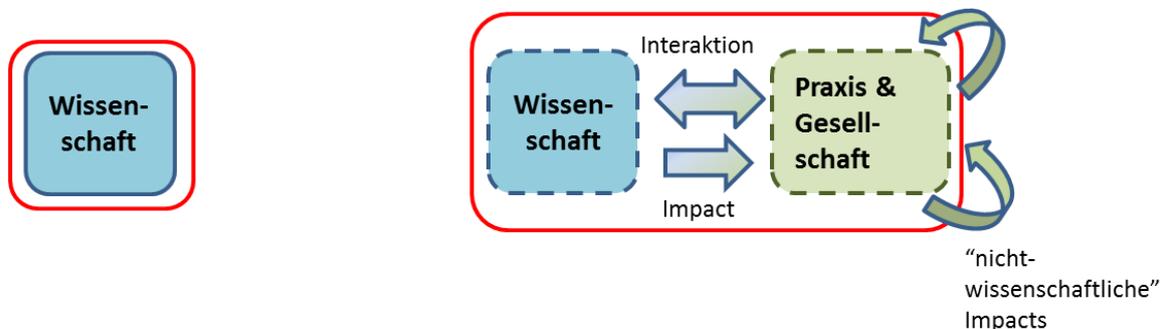


Abbildung 10: Evaluierung des wissenschaftlichen Impacts in einem wissenschaftsinternen, „geschlossenen“ System. Manuskript > Peer Review > Publikation > Zitation als linearer Prozess (links). Ebbvaluierung der Leistungen für Praxis und Gesellschaft in einem „offenen System“ mit wechselseitigen Interaktionen und außerwissenschaftlichen Ursachen für Impact.

6 Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse

Eine erste Verwertung der Ergebnisse erfolgte bereits, indem aus vorläufigen Projektergebnissen (Interviews und Literatur) auf Wunsch der BLE im September 2009 ein Arbeitspapier zur BÖL Evaluierung erstellt wurde. Dieses war Bestandteil des Ausschreibungsverfahrens der Evaluation der Forschung im BÖL.

Die veröffentlichten Ergebnisse des Projektes können von Dritten für eine verbesserte Evaluierung von Leistungen der Forschung für Praxis und Gesellschaft genutzt werden. Diese erarbeiteten Kriterien und Instrumente haben einen intensiven Diskussionsprozess mit Wissenschaftlern und der Forschungsförderung durchlaufen. Ihre genaue Auswahl müsste dabei anhand von Zielsetzungen, Evaluierungsgegenstand und vorhandenen Ressourcen fallspezifisch entschieden werden.

Eine vereinfachte Datenerfassung für die Evaluierungen von Leistungen der Forschung für Praxis und Gesellschaft ist eine wichtige Grundlage für die Etablierung einer solchen Evaluierung, da der bisherige Aufwand für die Evaluierung und dadurch auch die damit verbundenen Kosten erheblich reduziert werden können. Der Dokumentationsaufwand für die Wissenschaftler wird dabei nicht erhöht, wenn die bisherigen Berichte durch die strukturierte Dokumentation in Teilen ersetzt wird. Gleichzeitig sind durch eine vereinfachte Administration der Projekte auch Vorteile für die Projektträger zu erwarten. (Ausführlich in Kap. 5). Weil das entwickelte Konzept Vorzüge für alle beteiligten Akteure bietet, ist eine wichtige Grundlage für die zukünftige Akzeptanz und Anwendbarkeit der Evaluierung der Leistungen der Forschung für Praxis und Gesellschaft gelegt. Für eine Anwendung des Konzeptes in die-

sem Sinne, müsste es jedoch in einer Erprobung bis zu einem praxisreifen Dokumentationstool weiterentwickelt werden.

Für diese Erprobung liegt eine umfassende Arbeitsgrundlage vor: die detaillierte Auswertung der Interviewergebnisse (Anhang II) und das erste Konzept und Datenset für eine strukturierte Erfassung von Anträgen und Berichten (Anhang III). Eine Erprobung im Rahmen der Projektträger des BMELV wurde im Rahmen des Vortrags am 21.03.2011 mit der BLE und dem BMELV bereits grundsätzlich befürwortet. An einem Antrag für diese Erprobung wird gearbeitet, weshalb die im Anhang aufgeführten Ergebnisse nicht zur Veröffentlichung vorgesehen sind. Vorbehaltlich der Bewilligung eines solchen Antrags würde somit in ca. 3 – 4 Jahren ein erprobter „Prototyp“ für eine strukturierte, datenbankgestützte Dokumentation für die BLE bzw. das BMELV zur Verfügung stehen.

Sollte sich der Ansatz einer strukturierten Dokumentation von Forschungsleistungen beim BÖLN bzw. insgesamt bei den Projektträgern des BMELV durchsetzen, wäre damit ein vorbildliches Beispiel geschaffen, wie die Beiträge der Forschung zur praktischen Lösung gesellschaftlicher Probleme in der Forschungsevaluierung berücksichtigt werden können.

Zum einen könnte dadurch systemorientierte, praxisbezogene und interdisziplinäre Forschung zum Ökologischen Landbau angemessen bewertet und ihr Stellenwert auch außerhalb des BÖLN und des BMELV gestärkt werden. Zum anderen könnten zukünftig durch ein auf Ebene des BÖLN bzw. BMELV erprobtes Dokumentations- und Evaluierungskonzept, auch andere Projektträger animiert werden Leistungen für Praxis und Gesellschaft angemessen zu honorieren, wodurch eine verstärkte Ausrichtung der Forschung auf gesellschaftliche Problemlagen erfolgen würde. Eine Ausrichtung, die gegenüber drängenden Problemen wie z.B. Klimawandel, Ernährungssicherung und der Verlust von Biodiversität dringend erforderlich ist.

7 Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen

Veränderungen in der Zielsetzung des Projektes wurden bereits in Kap. 1.3 dargestellt und erläutert.

8 Zusammenfassende Schlussfolgerungen und Perspektiven

Die Projektergebnisse lassen die folgenden Schlussfolgerungen zu:

- Eine verstärkte Ausrichtung der Forschung auf einen gesellschaftlichen Nutzen bzw. Beiträge für eine nachhaltige Entwicklung wird von wissenschaftlichen und zivilgesellschaftlichen Initiativen zunehmend gefordert, wozu auch eine über die bisherigen Performanzindikatoren hinausgehende Forschungsevaluierung gehört.
- Für die Bewertung praxisorientierter und inter- und transdisziplinärer Forschung liegen vielfältige Konzepte vor, die sich zwar in Schwerpunkten und Detaillierungsgrad der Kriterien unterscheiden, aber eine große Schnittmenge aufweisen bezüglich der Leistungen, die für die Evaluierung inter- und transdisziplinärer Forschung relevant sind: Relevante Fragestellungen und Ergebnisse, Zusammenarbeit und Austausch mit anderen Wissenschaftlern und Akteuren aus Praxis und Gesellschaft, sowie Publikationen und Produkte, Relevanz und tatsächlicher Impact für Praxis und Gesellschaft.
- Forschungsleistungen für Praxis und Gesellschaft werden überwiegend mittels aufwändiger Verfahren wie Dokumentenanalysen und Interviews erfasst, weil Evaluierungen bisher als „einzelne Vorhaben“ durchgeführt werden, wodurch die Daten auch nicht für andere Evaluierungen verfügbar gemacht werden.
- Es liegen kaum erprobten Konzepte vor, wie der „Scientific Impact“ grundsätzlich um eine Evaluierung der Leistungen von Forschung für Praxis und Gesellschaft ergänzt werden kann. Erste Konzepte sind die initiative „Excellence with Impact“ der Research Councils UK und das Standard Evaluation Protokoll (SEP) der Niederländischen Universitäten.
- Für die Entwicklung einer solchen grundsätzlichen Ergänzung des „Scientific Impact“ wird es als notwendig angesehen, möglichst keinen zusätzlichen Aufwand für die Datenerhebung zu bedingen und die Daten für verschiedene Evaluierungsanlässe nutzbar zu machen. Das hier vorgestellte Konzept sieht deshalb vor, derzeit im Rahmen von Projekten unumgängliche Dokumentationsleistungen (z. B. Anträge und Berichte) stärker zu strukturieren und formalisieren und damit auch eine Datengrundlage für die Evaluierung zu schaffen. Die Speicherung und Verwaltung der Eingaben könnte in einem Datenbanksystem erfolgen, das sowohl für die Projektverwaltung bei Forschungsförderungseinrichtungen als auch für verschiedene Evaluierungsverfahren genutzt wird.
- Wichtige Entwicklungen zur Standardisierung der Datenerfassung zeigen sich im Bereich des Hochschul-Rankings, da hier einzelne Aspekte wie Wissenstransfer oder gesellschaftlicher Nutzen zunehmend in Betracht gezogen werden. Da die Dokumentation hierfür auf institutioneller Ebene erfolgt, zeigten sich jedoch Defizite in der Datenverfügbarkeit. Das in diesem Projekt entwickelte Konzept würde es ermöglichen, die Datenbasis für Leistungen der Forschung für Praxis und Gesellschaft grundlegend zu verbessern und auch für Institutionen verfügbar zu machen.

- Ob bzw. wie das vorgestellte Konzept konkret umgesetzt werden kann und in welchem Ausmaß der erwartete Nutzen für Evaluation, Forschungsförderung und Wissenschaftler erzielt werden kann, muss weiter erforscht werden, z. B. durch Entwicklung und Erprobung anhand von Fallstudien.
- Die Entwicklung und Erprobung der strukturierten Dokumentation innerhalb von Fallstudien sind durch einen Diskurs zu begleiten, in dem laufend Ergebnisse kritisch reflektiert werden und die Bedeutung einzelner Kriterien für verschiedene Evaluierungsgegenstände bewertet wird. Für diese national und international bereits angestoßene Entwicklung sind nicht nur Vertreter und Entscheidungsträger aus der Forschung und Forschungsförderung sondern auch Akteure aus Praxis und Gesellschaft einzubeziehen, um auch deren Sichtweisen und Anforderungen angemessen zu integrieren.
- Hinsichtlich der Instrumente für die Einbindung von Akteuren aus Praxis und Gesellschaft sollte sowohl eine Integration von Feedback in das Datenbanksystem, als auch die Entwicklung dauerhafter, verbindlicher, repräsentativer, fachübergreifender Dialogstrukturen im Fokus stehen.
- Letztlich zielt die Berücksichtigung der Wirkungen von Forschung auf Praxis und Gesellschaft in der Forschungsevaluierung darauf ab, dass die Forschung ihre Verantwortung gegenüber der Gesellschaft in ausreichendem Maße wahrnimmt und bezüglich ihrer Wirkungen, wie auch ihrer Finanzierung einer demokratischen Kontrolle zugänglich wird (vgl. VDW 2010a).

9 Literaturverzeichnis

- Aeberhard, A., und S. Rist. 2009. Transdisziplinäre Wissensproduktion im Biolandbau. *AGRARForschung* 16, no. 10:414–19.
http://www.agrarforschungschweiz.admin.ch/archiv_11de.php?id_artikel=1514 (accessed September 28, 2010).
- Aennis, T., und U. J. Nagel. 2003. Impact indicator definition within a transdisciplinary Forschungsgruppe. In *Interdisciplinary and transdisciplinary landscape studies: Potential and limitations*, ed. Bärbel Tress, Gunther Tress, Arnold van der Valk, und Gary Fry, 160–69. Delta series 2. Wageningen: Delta Programm.
- Aföldi, T., U. Niggli, S. Bellon, und J. Blanc. 2007. *Coordination of European Transnational Research in Organic Food and Farming; Projekt no. ERAC-CT-2004-011716*.
- Albrecht, S., G. Feuerstein, I. Schneider, und S. Stirn. 2008. *Konstitution und Evaluation von inter- und transdisziplinärer Forschung unter besonderer Berücksichtigung der Anforderungen von Technikfolgenabschätzung und -bewertung*.
- Atteslander. 2000. *Methoden der empirischen Sozialforschung*. 9th ed. Berlin: Walter de Gruyter.
- AUBR EG. 2009. *Assessing Europe's University-based research: Expert group o assessment of University-based research*. http://ec.europa.eu/research/science-society/document_library/pdf_06/assessing-europe-university-based-research_en.pdf (accessed July 4, 2011).
- Begusch-Pfefferkorn, K. 2006. KLF-Wissensbilanz – Leistungsdarstellung des inter- und transdisziplinären Programms Kulturlandschaftsforschung. In *Fachübergreifende Nachhaltigkeitsforschung – Stand und Visionen am Beispiel nationaler und internationaler Forscherverbände*, ed. Bernhard Glaeser, 119–39. Edition Humanökologie. Munich: oekom.
- Beisiegel, U., Burkhardt, A., Dichgans, J., Felt, U., Frühwald, W., Götz, M., Hornbostel, S., Jacobs, R., Jahn, R., Kempermann, G., Konze-Thomas, B., Leitner, M., May, T., Mihr, A., Pulverer, B., Schnabel, U., von Figura, K., Weingart, P., Wolfrum, R., Wormer, H. 2011. *Wie die Wissenschaft Integrität und Qualität sichern kann - Thesenpapier der Gäste des 4. Berliner Wissenschaftsgesprächs der Robert Bosch Stiftung*. http://www.bosch-stiftung.de/content/language1/downloads/Thesenpapier_BWG_4.pdf
- Bergmann, M., B. Brohmann, E. Hoffmann, M. C. Loibl, R. Rehaag, E. Schramm, und J.-P. Voß. 2005. *Qualitätskriterien transdisziplinärer Forschung: Ein Leitfaden für die formative Evaluation von Forschungsprojekten*. Frankfurt am Main.
- Bergmann, M., T. Jahn, T. Knobloch, W. Krohn, C. Pohl, und E. Schramm. 2010. *Methoden transdisziplinärer Forschung*. Frankfurt am Main: Campus Verlag.
- Blackstock, K. L., G. J. Kelly, und B. L. Horsey. 2007. Developing and applying a framework to evaluate participatory research for sustainability. *Ecological Economics*, no. 60:726–42.
- BLE. *Richtlinien für Zuwendungsanträge auf Ausgabenbasis*. http://www.kp.dlr.de/profi/easy/formular_ble.html.

- BMBF (Hg.) 2009. Forschung für nachhaltige Entwicklungen. Rahmenprogramm des BMBF. Berlin.
http://www.bmbf.de/pubRD/leitlinien_ressortforschung.pdf.
- BMBF. 2006. *Besondere Nebenbestimmungen für Zuwendungen des Bundesministeriums für Bildung und Forschung zur Projektförderung auf Ausgabenbasis (BNBest-BMBF 98)*.
- BMELV 2008. Förderprogramm Nachwachsende Rohstoffe Programm des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz.
http://www.nachwachsenderohstoffe.de/fileadmin/fnr/pdf/Brosch_Foerderprogramm7_BMELV.pdf.
- BMELV 2009. Programm zur Innovationsförderung.
http://www.bmelv.de/SharedDocs/Downloads/Broschueren/Innovationsfoerderung.pdf?__blob=publicationFile.
- BMELV 2010. Programm des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz zur Förderung von Forschungs- und Entwicklungsvorhaben sowie von Maßnahmen zum Technologie- und Wissenstransfer im ökologischen Landbau.
http://www.bundesprogramm.de/fileadmin/sites/default/files/Forschung/Richtlinie_Forschungsforderung_2013.pdf.
- BMU 2011. Richtlinien des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit zur Förderung von Maßnahmen im Rahmen des Bundesprogramms Biologische Vielfalt vom 26. Januar 2011.
http://www.biologischerdiversitaet.de/fileadmin/NBS/documents/Bundesprogramm/FoeRiLi/BPBV_FoeRiLi_Abgestimmt_neu_mit_Logo_20110211.pdf.
- Böcher, M., und M. Krott. 2010. *Umsetzung des Konzepts einer modernen Ressortforschung im Geschäftsbereich des BMU*. im Auftrag des Umweltbundesamtes. Dessau-Roßlau.
<http://www.uba.de/uba-info-medien/3995.html>.
- Boix Mansilla, V., und H. Gardner. 2003. *Assessing interdisciplinary work at the frontier: an empirical exploration of 'symptoms of quality'*.
- BÖLN 2011: <http://www.bundesprogramm.de/index.php>, <http://www.bundesprogramm.de/das-programm/inhalt-und-struktur/9>, <http://www.bundesprogramm.de/forschungsmanagement/>.
- Brown, H. 2007. *How impact factors changed medical publishing - and science*. BMJ, 334, 561-564.
- CHERPA-Network. 2010. *U-Multirank - Project "Design and Testing the Feasibility of a Multi-dimensional Global University Ranking": Interim progress report: Preparation of the pilot phase*.
<http://www.u-multirank.eu/project/U-Multirank%20Interim%20Report%202.pdf>.
- Clark, W. C., P. J. Crutzen, und H. J. Schellnhuber. 2005. *Science for Global Sustainability: Toward a New Paradigm*. CID Working Paper No. 120, Harvard University.
- Defila, R., A. Di Giulio, und M. Scheuermann. 2008. *Management von Forschungsverbänden – Möglichkeiten der Professionalisierung und Unterstützung*. Reihe Standpunkte. Weinheim.
- Defila, R., und A. Di Giulio. 1999. Evaluationskriterien für inter- und transdisziplinäre Forschung. *Panorama*, no. 99:5–11. <http://www.ika.oe.unibe.ch/forschung/ip/Sondernummer.Pano.1.99.pdf>.

- DeGEval. 2008. *Standards für Evaluation*. 4th ed. Hamburg: Medienzirkus Gudrun Schwank.
- EC. 2006. *REGULATION (EC) No 1906/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 laying down the rules for the participation of undertakings, research centres and universities in actions under the Seventh Framework Programme and for the dissemination of research results (2007-2013)*. Brüssel. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:391:0001:0018:EN:PDF> (accessed May 30, 2011).
- Elbe, S. 2007. *Die Voraussetzungen der erfolgreichen Steuerung integrierter Ansätze durch Förderprogramme: Untersucht am Beispiel des Modellvorhabens Regionen Aktiv*. Universität Dortmund.
- ERIC 2010: Evaluation the societal relevance of academic research: A guide.
- Federkeil, G. 2011. U-Multirank - Ein anderes internationales Ranking ist möglich! *Checkup*, no. 1. http://www.che-ranking.de/downloads/checkup/checkup_50.pdf (accessed July 8, 2011).
- Flick, U. 2002. *Qualitative Forschung, ein Handbauch*. 2nd ed. Reinbek, Hamburg: Rowohlt Taschenbuch Verlag.
- Formas. 2009. *Evaluation report 2009. Evaluation of SLF-funded Research 2002–2007*. http://www.formas.se/upload/EPiStorePDF/Evaluation_of_SLF_funded_research_2002_2007_%20R3_2009/Formas_SLF.pdf.
- Frank-Rieser, E. 1999. *Leistungskriterien, Konkurrenz und Teamarbeit*. BUKO Info 4/99, 23-28. Wien.
- Frodeman, R., und J. B. Holbrook. 2007. Science's Social Effects. *Issues in Science and Technology*:28–30.
- Fröhlich, G. 1999. *Das Messen des leicht Messbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie?*. GMD-Report 61, 27-38.
- Fröhlich, G. 2003a. Evaluation wissenschaftlicher Leistungen. In *Medizin-Bibliothek-Information*, 3 (2), 29-32.
- Fröhlich, G. 2003b. *Der Impact-Faktor auf dem Prüfstand der Wissenschaftsforschung*. BUKO Info, 1-4, 61-65.
- Gerber, A., V. Hoffmann, und A. Thomas. 2009. *Transdisziplinäre Umweltforschung: Methodenhandbuch Kulturlandschaft*. 2nd ed. Munich: oekom.
- Glaeser, B. 2007. Anforderung an eine Evaluation der inter- und transdisziplinären Humanökologie und Nachhaltigkeitsforschung aus Sicht der Humanökologie. In *Evaluation inter- und transdisziplinärer Forschung: Humanökologie und Nachhaltigkeitsforschung auf dem Prüfstand*, ed. Susanne Stoll-Kleemann und Christian Pohl, 85–93. Edition Humanökologie. Munich: oekom.
- Hennen, L., T. Petermann, und C. Scherz. 2004. *Partizipative Verfahren der Technikfolgen - Abschätzung und parlamentarische Politikberatung*. Neue Formen der Kommunikation zwischen Wissenschaft, Politik und Öffentlichkeit.
- Herb, U. 2008. Vermessung der Wissenschaft. In *Telepolis*: <http://www.heise.de/tp/r4/artikel/28/28942/1.html>. Heise Zeitschriften Verlag.

- Hornbostel, S. 1999. Evaluation und Ranking – Führen sie zu mehr Transparenz und Vergleichbarkeit? In *Hochschulrektorenkonferenz 1999* (Hg.). „Viel Lärm um Nichts?“ *Evaluation von Studium und Lehre und ihre Folgen*. Tagung Univ. Rostock, 6.-8. Sept. 1998, Beiträge zur Hochschulpolitik 4/1999, Berlin
- IAASTD 2009. *Weltagrarbericht: Synthesebericht*. http://hup.sub.uni-hamburg.de/opus/volltexte/2009/94/pdf/HamburgUP_IAASTD_Synthesebericht.pdf (accessed December 17, 2010).
- IFOAM. 2008. *The Principles of Organic Agriculture*. http://www.ifoam.org/about_ifoam/principles/index.html.
- Imboden, D. (2006): *Evaluation von Wissenschaft oder Die verlorene Unschuld*. GAIA 12, 3, 167-168.
- Klein, J. T. 2006. Afterword: the emergent literature on interdisciplinary and transdisciplinary research evaluation. *Research Evaluation* 15, no. 1:75–80.
- Krott, M. 2004. Evaluation of transdisziplinäre research. In *Unity of Knowledge (in Transdisciplinary Research for Sustainability)*, ed. Gertrude Hirsch Hadorn. Encyclopedia of Life Support Systems (EOLSS) Developed under the Auspices of the UNESCO Eolss Publishers OxfordUK. [<http://www.eolss.net> (accessed October 30, 2010)].
- Kueffer, C., Hirsch-Hadorn, G., Bammer, G. Van Kerkhoff, L., Pohl, C. (2007): Towards a Publication Culture in Transdisciplinary Research. GAIA 16/1, 22–26.
- Lamnek S. 1995. *Qualitative Sozialforschung*. Band 1. Methodologie. 3rd ed. Weinheim: Psychologie Verlags Union.
- Lampkin, N. 1990. *Organic Farming*. Ipswich: Farming Press Books.
- Lindenthal, T., C. R. Vogl, und J. Heß. 1996. Forschung im Ökologischen Landbau: Integrale Schwerpunktthemen und Methodikkriterien. *Sonderausgabe der Zeitschrift Förderungsdienst*, no. 2c.
- Lyall, C., A. Bruce, J. Firn, M. Firn, und J. Tait. 2004. Assessing end-use relevance of public sector research organisations. *Research Policy* 33:73–87.
- Neidhardt, F. 2010. Selbststeuerung der Wissenschaft: Peer Review. In Simon, D., Knie, A., Hornbostel, S. (Hg.): *Handbuch Wissenschaftspolitik*. Wiesbaden (VS Verlag für Sozialwissenschaften) 280-293.
- Niggli, U., und A. Gerber. 2010. *Öko-Forschung: Innovationsmotor für eine zukunftsfähige Landwirtschaft*. Berlin.
- NSF. 2007. *Merit Review Broader Impacts Criterion: Representative Activities*. <http://www.nsf.gov/pubs/gpg/broaderimpacts.pdf> (accessed May 27, 2011).
- Martin, B. R. (2011): The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? In: *Research Evaluation* 20 (3), S. 247–254.
- Pedersen, S. M., M. V. Boesen, D. Baker, und A. Larsen 2009. A model for evaluating Food and Agricultural Research Projekts: A Danish application. In *Proceedings of the 17th International Farm*

- Management Congress, Bloomington/Normal, Illinois, USA: Agriculture: Food, Fiber and Energy for the Future*”, ed. Harold H. Guither, Jean L. Merry, und Carroll E. Merry, 99–113.
- Pietschmann, H. (1980a): *Das Ende des naturwissenschaftlichen Zeitalters*. Zsolnay-Verlag. Wien.
- Pietschmann, H. (1980b): *Wissenschaftstheoretischer Vortrag an der Universität für Bodenkultur*, Wien (unveröff./Tonband-Aufnahme).
- Pohl, C., und G. Hirsch Hadorn. 2006. *Gestaltungsprinzipien für die transdisziplinäre Forschung*. Munich: oekom.
- Power, M. (1997): *The Audit Society. Ritual of Verification*, Oxford University Press, Oxford.
- RCUK (Research Councils UK). 2010. Research outcomes project invitation to tender research outcome types. Swindon, UK: RCUK. www.rcuk.ac.uk/documents/oocp/OutputTypes.pdf (Stand: September 21, 2011).
- RCUK. 2011a. RCUK impact requirements: Frequently asked questions. Swindon, UK: RCUK. www.rcuk.ac.uk/documents/impacts/RCUKImpactFAQ.pdf (Stand: September 21, 2011)
- RCUK. 2011b. RCUK research outcomes system now available. Swindon, UK: RCUK. www.rcuk.ac.uk/media/news/2011news/Pages/111124.aspx (accessed October 24, 2012).
- Regulation (EC)No1906/2006. 2006.
- Reuber, M., und O. Haas. 2009. *Evaluierungen in der Deutschen Gesellschaft für Technische Zusammenarbeit (GTZ)*.
- Schavan, A., ed. 2008. *Keine Wissenschaft für sich: Essays zur gesellschaftlichen Relevanz von Forschung*. Hamburg: Ed. Körber-Stiftung.
- Schiller, J., R. Manstetten, B. Klauer, P. Steuer, H. Unnerstall, H. Wittmer, und B. Hansjürgens. 2005. *Evaluation interdisziplinärer Programm-Forschung - Ziele, Anlässe und Prozesse: UFZ-Diskussionspapiere 17/2005*. UFZ-Diskussionspapiere. Leipzig.
- Schlinghoff, A. und U. Backes-Gellner 2002. *Publikationsindikatoren und die Stabilität wirtschaftswissenschaftlicher Zeitschriftenrankings*. Zeitschrift für betriebswirtschaftliche Forschung (zfbf), 54, Juni/2002, 343-362.
- Schmid O., D. S., Eichert C., González V., Lampkin N., Michelsen J., Slabe A., Stokkers R. et al. 2008. *Organic Action Plans*. <http://orgprints.org/13481>.
- Schmid, O., S. Padel, N. Halberg, M. Huber, I. Darnhofer, C. Micheloni, Koopmans C. et al. 2009. *Strategic Research Agenda for organic food and farming*. Brüssel.
- Schübel, H. R. 2007. Interdisziplinäre Kooperation in transdisziplinärer Forschung besser machen? Prozessbegleitung als formative Evaluation. In *Evaluation inter- und transdisziplinärer Forschung: Humanökologie und Nachhaltigkeitsforschung auf dem Prüfstand*, ed. Susanne Stoll-Kleemann und Christian Pohl, 195–210. Edition Humanökologie. Munich: oekom.
- SIAMPI. 2010. *SIAMPI approach for the assessment of social impact: Summary for participants*.
- Smoliner, C. 2007. Evaluation in der Programmforschung oder doch lieber würfeln? In *Evaluation inter- und transdisziplinärer Forschung: Humanökologie und Nachhaltigkeitsforschung auf dem Prüfstand*, ed. Susanne Stoll-Kleemann und Christian Pohl, 95–108. Edition Humanökologie. Munich: oekom.

- Spaapen, J., H. Dijstelbloem, und F. Wamelink. 2007. *Evaluation research in context: A method for comprehensive assessment*. The Hague.
- Stock, W.G. 1994. *Wissenschaftsevaluation. Die Bewertung wissenschaftlicher Forschung und Lehre*. Ifo-Diskussionsbeiträge, Nr. 17, S. 46-51, München.
- Stockmann, R. 2004. *Was ist eine gute Evaluation: Einführung zu Funktionen und Methoden von Evaluationsverfahren*. <http://www.ceval.de> (accessed September 2, 2010).
- Stokols, D., J. Fuqua, J. Gress, R. Harvey, K. Phillips, L. Baezconde-Garbanati, J. Unger et al. 2003. Evaluating transdisciplinary science. *Nicotine & Tobacco Research*, no. 5:21–39.
- Stoll-Kleemann, S. 2007. Potentiale der Evaluation inter- und transdisziplinärer Nachhaltigkeitsforschung und Humanökologie. In *Evaluation inter- und transdisziplinärer Forschung: Humanökologie und Nachhaltigkeitsforschung auf dem Prüfstand*, ed. Susanne Stoll-Kleemann und Christian Pohl, 25–40. Edition Humanökologie. München: oekom.
- VDW. 2010. *Für eine verantwortbare und zukunftsorientierte Forschungspolitik in Deutschland: VDW Positionspapier*. http://www.vdw-ev.de/images/stories/vdwdokumente/aktuelles/vdw_papier_zur_forschungspolitik.pdf (accessed May 16, 2011).
- VDW. 2010a. *Für eine verantwortbare und zukunftsorientierte Forschungspolitik in Deutschland: VDW Positionspapier*. http://www.vdw-ev.de/images/stories/vdwdokumente/aktuelles/vdw_papier_zur_forschungspolitik.pdf (accessed May 16, 2011).
- VisionRD4SD. 2011. *The Challenge of Research and Development for Sustainable Development (RD4SD)*. <http://www.visionrd4sd.eu/top-about-vision-rd4sd/rationale> (accessed July 7, 2011).
- Walter, A. I., S. Helgenberger, A. Wiek, und Scholz R. W. 2007. Measuring societal effects of transdisciplinary research Projekts: Design and application of an evaluation method. *Evaluation and Programm Planning*, no. 30:325–38.
- WBGU. 2011. *Welt im Wandel Gesellschaftsvertrag für eine große Transformation*. Berlin. http://www.wbgu.de/fileadmin/templates/dateien/veroeffentlichungen/hauptgutachten/jg2011/wbgu_jg2011.pdf (accessed July 6, 2011).
- Wissenschaftsrat. 2007. *Kriterien des Ausschusses Ressortforschung für die Begutachtung von Bundeseinrichtungen mit FuE-Aufgaben*. Köln.
- Wissenschaftsrat. 2009. *Bewertungsmatrix für das Forschungsrating in der Elektro- und Informationstechnik: Fassung vom 11.12.2009 mit Fragebogenbezügen*. <http://www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Bewertungsmatrix-ETIT.pdf>.
- Wissenschaftsrat. 2010. *Arbeitsprogramm 2010/2011*.
- Young, N. S., Ioannidis, J. P. A., Al-Ubaydli, O. (2008): *Why Current Publication Practices May Distort Science*. In: Public Library of Science PLoS, Medicine, Vol. 5, Issue 10, e201, 1-5.

10 Wissenstransfer

10.1 Erfolgte und geplante Veröffentlichungen

Wolf B., Lindenthal T., Szerencsits M., Heß J. (2011): Praxisorientierte Forschung. Was muss sich in der Evaluierung verändern? *Ökologie und Landbau* 157, 2011/1, S. 54-56.

Wolf B., Lindenthal T., Szerencsits M., Boland H., Heß J. (2011): Evaluierung praxisorientierter Ökolandbauforschung. Beitrag für die Wissenschaftstagung zum Ökologischen Landbau 2011. Verlag Dr. Köster, Band 2, S. 336-339. bzw. http://orgprints.org/17714/3/Wolf_17714.pdf.

Manuskripte in 09/2011 kurz vor der Einreichung:

Wolf, B., Lindenthal, T., Szerencsits, M., Heß, J. Evaluation of science regarding its impact on practice and society - suggestions for a complementary assessment from the perspective of practice-oriented research on organic agriculture.

Lindenthal, T., Wolf, B., Heß, J. Critical analysis of established evaluation criteria and instruments from the point of view of organic agriculture and sustainability research

10.2 Aktivitäten zur Verbreitung der Ergebnisse

Beiträge zu Konferenzen:

Vortrag: „Evaluation praxisorientierter Ökolandbauforschung“ Vortrag zur Wissenschaftstagung zum Ökologischen Landbau, Gießen 16.03.2011 (Birge Wolf, Manfred Szerencsits)

Workshop (17.03.2011) „Bewirkt Forschung etwas für die praktische und konsequent nachhaltige Entwicklung des Biolandbaus?“ Impulsreferat (Birge Wolf) „Diskussion zur Evaluierung praxisorientierter Ökolandbauforschung – Forscherglück braucht Bauernblick?!“ (Präsentation spezifisch für Wissenschaftler und Praktiker vorbereitet, die für Wissenschaftler wurde gehalten, da keine Praktiker anwesend waren.) In: Brock C., Wilbois, K.-P., Becker K., Fischinger S. Spiegel A.-K., Spory K., Williges U., Zerger U., Leithold G. (Hrsg.) (2011): Es geht ums Ganze: Forschen im Dialog von Wissenschaft und Praxis. Ergebnisse der Dialogworkshops bei der 11. Wissenschaftstagung Ökologischer Landbau (15.-17.März 2011 in Gießen). http://orgprints.org/19068/1/WiTa2011_DWDoku.pdf, S. 53

Workshop (17.03.2011): „Berater-Praxis-Netzwerke: Chancen, Probleme, Perspektiven praxisorientierter Netzwerkarbeit im Ökologischen Landbau“ Impulsreferat (Jürgen Heß): „Perspektiven praxisorientierter Netzwerkarbeit aus Sicht der Wissenschaft“. In: Brock C., Wilbois, K.-P., Becker K., Fischinger S. Spiegel A.-K., Spory K., Williges U., Zerger U., Leithold G. (Hrsg.) (2011): Es geht ums Ganze: Forschen im Dialog von Wissenschaft und Praxis. Ergebnisse der Dialogworkshops bei der 11. Wissenschaftstagung Ökologischer Landbau (15.-17.März 2011 in Gießen). http://orgprints.org/19068/1/WiTa2011_DWDoku.pdf S.44

Vorträge an Universitäten und Instituten:

- Bodenkundliche Universität Wien, Arbeitsgruppe Nachhaltigkeitsbewertung 18.06.2010 (Wissenschaftler)
- Universität Kassel-Witzenhausen, FB 11, 30.06.2010, (Wissenschaftler)
- Landesbetrieb Landwirtschaft Hessen, Alsfeld, 07.09.2010, (Berater)
- Gießen, Arbeitsgruppen Prof. Boland (Institut für Agrarsoziologie und Beratungswesen) und Prof. Leithold (ökologischer Landbau), 16.11.2010, (Wissenschaftler)
- BIOGUM Hamburg, Forschungsschwerpunkt Landwirtschaft 19.10.2011 (Wissenschaftler)
- Trenthorst, Institut für Ökologischen Landbau (vTI), 07.12.2010 (Wissenschaftler und weitere Mitarbeiter des Instituts)

Geplante Aktivitäten:

Vortrag *“Establishing Complementary Assessment of Science Regarding Its Impact on Practice and Society as an Integral Element of Research Evaluation”* Beitrag zur Konferenz „Evaluation of transdisciplinary Research“ Konferenz des td-net (14. – 16.09.2011, Bern)
<http://www.transdisciplinarity.ch/e/Conference/international/2011/index.php>.