

Lars Balzer

Die Evaluationspraxis im Licht der neuen SEVAL-Standards sowie empirisch bestimmter Bedingungen erfolgreicher Evaluationsprojekte

Dieser Beitrag berichtet von zwei empirischen Studien, die Erfolgsbedingungen von Evaluationsprojekten empirisch untersuchen und diese in Bezug zu den neuen SEVAL-Standards setzen. Dabei zeigt sich unter anderem, dass die neuen SEVAL-Standards ziemlich gut dem entsprechen, was Evaluationsexpertinnen und -experten als Bedingungen erfolgreicher Evaluationsprojekte ansehen. Gleichzeitig werden in der aktuellen Evaluationspraxis Defizite festgestellt, was die Grundlagen konkreten Handelns in Evaluationsprojekten betrifft, und entsprechender Handlungsbedarf identifiziert.

Beitragsart: Wissenschaftliche Beiträge

Zitiervorschlag: Lars Balzer, Die Evaluationspraxis im Licht der neuen SEVAL-Standards sowie empirisch bestimmter Bedingungen erfolgreicher Evaluationsprojekte, in: LeGes 30 (2019) 3

Inhaltsübersicht

1. Einstieg
2. SEVAL-Standards
3. Studien zur Evaluationspraxis
 - 3.1. Forschungsfragen
 - 3.2. Durchführung
 - 3.3. Realisierte Stichproben
 - 3.4. Ergebnisse: Erfolgsbedingungen gestern und heute
 - 3.5. Ergebnisse: Erfolgsbedingungen und Evaluationsstandards
 - 3.6. Ergebnisse: Handlungsbedarf in der Evaluationspraxis
 - 3.7. Methodische Anmerkungen
4. Einordnung in die Diskussion zur Professionalisierung der Evaluation in der Schweiz

1. Einstieg

[1] Evaluation spielt in der Schweiz eine wichtige Rolle.

[2] Sie hat sogar Verfassungsrang. Seit der Totalrevision der Bundesverfassung (BV) im Jahr 1999 heisst es in Artikel 170 BV: «Die Bundesversammlung sorgt dafür, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden».¹ Im französischen Wortlaut wird statt Wirksamkeitsüberprüfung der Begriff Evaluation sogar direkt verwendet: «L'Assemblée fédérale veille à ce que l'efficacité des mesures prises par la Confédération fasse l'objet d'une évaluation».² Was darunter zu verstehen ist, wurde vielerorts ausführlich beschrieben. MADER (2005) fasst zusammen, dass damit Regierung und Verwaltung in die Pflicht genommen werden, dass staatliche Aktivitäten sehr umfassend gemeint sind, dass nicht nur nachträgliche, sondern auch prospektive Überprüfungen eingeschlossen sind und dass neben der Wirksamkeit im Sinne von Zielerreichungsgrad generell die Eruiierung und Beurteilung der Wirkungen von Massnahmen gemeint sind, ob also z. B. «bestimmte Massnahmen entscheidungskonform umgesetzt werden oder umgesetzt werden können (Effektivität, Befolgung, Durchsetzung, Inanspruchnahme) und ob sie neben der Erreichung der gesetzten Ziele noch andere Auswirkungen haben (Nebenwirkungen, Spätfolgen usw.)» (MADER 2005, 30f).

[3] Der Schweiz wird von verschiedener Seite eine fortgeschrittene Evaluationskultur bescheinigt (z. B. BALTHASAR 2007, HORBER-PAPAZIAN 2015). SAGER, WIDMER und BALTHASAR (2017) sprechen von einem guten Nährboden für die Entwicklung einer Evaluationskultur, zumindest in neuerer Zeit. Auch fragen Parlamentarierinnen und Parlamentarier in der Schweiz vermehrt Evaluationen nach (BUSSMANN 2010). Folgt man JACOB, SPEER und FURUBO (2015), so hat die Schweiz nach Finnland sogar die am höchsten entwickelte Evaluationskultur von 19 untersuchten OECD-Ländern, was sich u. a. in einer am weitesten entwickelten Institutionalisierung der Evaluation in den Parlamenten und den politischen Systemen zeigt. Deren Entwicklung bis 2008 beschreibt BUSSMANN (2008) ausführlich. Aktuellere Ausführungen zu deren Besonderheiten und Umsetzungen finden sich in HORBER-PAPAZIAN (2015).

[4] Einen anderen Blick auf die Evaluation in der Schweiz werfen SAGER, WIDMER und BALTHASAR (2017). Mit Ergebnissen aus verschiedenen Studien wird u. a. beschrieben, welche Bedeutung typische Charakteristiken des schweizerischen politischen Systems – wie Föderalis-

¹ <https://www.admin.ch/opc/de/classified-compilation/19995395/index.html>.

² <https://www.admin.ch/opc/fr/classified-compilation/19995395/index.html>.

mus oder direkte Demokratie – für die Entwicklung der Evaluation in der Schweiz haben, wie Evaluationen die Problemlösungsfähigkeit des schweizerischen politischen Systems beeinflussen, welchen Beitrag Evaluationen zur Stärkung der Demokratie und zur Versachlichung des politischen Diskurses leisten und ob Evaluationen weiterhin eine Berechtigung im politischen System der Schweiz haben.

[5] Insgesamt ist in der Schweiz über die Jahre eine steigende Akzeptanz von Evaluation zu beobachten (BUSSMANN 2015). «Die Evaluation, verstanden als wissenschaftliche Dienstleistung zur systematischen und transparenten Bewertung eines Gegenstandes, hat sich in den letzten beiden Dekaden in der Schweiz in vielen Feldern öffentlicher Politik etabliert» (WIDMER 2011, 24). WIDMER, BEYWL und FABIAN (2009) oder auch WIDMER und DE ROCCHI (2012) liefern hierfür einen detailreichen Überblick.

[6] Für diese steigende Akzeptanz auf dem Hintergrund des genannten verfassungsrechtlichen Rahmens wurde viel unternommen. Eine dieser zentralen Aktivitäten betrifft die Entwicklung und Etablierung von Evaluationsstandards.

[7] Der vorliegende Beitrag stellt diese unter verschiedenen Perspektiven ins Zentrum. Dazu wird im nachfolgenden Kapitel 2 zunächst beschrieben, was Evaluationsstandards sind, wie sie in der Schweiz entwickelt wurden und wie sie in der Evaluationspraxis Verwendung finden. Daraus abgeleitete Forschungsfragen leiten das Kapitel 3 ein, das sich einer Studie zur Evaluationspraxis widmet und zunächst die Durchführung der empirischen Arbeit sowie die realisierten Stichproben beschreibt. Daraufhin folgen empirische Ergebnisse zu Erfolgsbedingungen von Evaluationen, zu deren Bezug zu Evaluationsstandards sowie zu aktuellem Handlungsbedarf für die Evaluationspraxis. Mögliche methodische Einschränkungen der Studie werden reflektiert, bevor im abschliessenden Kapitel 4 versucht wird, die Ergebnisse in die aktuelle Diskussion zur Professionalisierung der Evaluation in der Schweiz einzuordnen³.

2. SEVAL-Standards

[8] Evaluationsstandards wurden in der Schweiz von der Schweizerischen Evaluationsgesellschaft SEVAL⁴ entwickelt. Die erste Version wurde 2001 vorgelegt (WIDMER/LANDERT/BACHMANN 2001), ein bei RÜEFLI (2013) sowie RÜEFLI und ZWEIDLER (2019) beschriebener Revisionsprozess führte 2016 zur zweiten Version (SEVAL – Schweizerische Evaluationsgesellschaft 2016). Die SEVAL-Standards definieren Regeln professioneller Evaluation und formulieren Qualitätsanforderungen an gute Evaluationen. Zu Zweck und Geltungsbereich heisst es in deren Einführungstext auf Seite 1: «Die Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards) leisten einen Beitrag zur Professionalisierung der Evaluationstätigkeit in der Schweiz. Sie formulieren zentrale Prinzipien, deren Beachtung die Qualität und Glaubwürdigkeit von Evaluationen erhöht. Die SEVAL-Standards bieten eine Grundlage und Hilfestellung für Planung und Durchführung von Evaluationen, für die Formulierung von Aufträgen, Ausschrei-

³ Der Autor bedankt sich herzlich bei Christian Rüefli für sein konstruktiv-kritisches Feedback zu einer früheren Fassung dieses Textes.

⁴ Die Schweizerische Evaluationsgesellschaft SEVAL wurde 1996 mit dem Ziel gegründet, den Dialog sowie den Informations- und Erfahrungsaustausch zwischen Politik, Verwaltung, Hochschulen und Beratungsfirmen zu fördern. Ausserdem soll die Qualität von Evaluationen verbessert und die Verwendung von Evaluationen verbreitet werden, vgl. <https://www.seval.ch/ueber-uns/>.

bungen und Evaluationsvereinbarungen und für die begleitende Sicherung oder nachträgliche Beurteilung der Qualität von Evaluationen. Sie dienen auch als didaktisches Hilfsmittel für die professionelle Aus- und Weiterbildung. Die SEVAL-Standards sind grundsätzlich für alle Arten von Evaluationen anwendbar, unabhängig vom institutionellen Kontext, dem gewählten Vorgehen und dem spezifischen Themenbereich sowie unabhängig davon, ob es sich um eine externe, interne oder Selbstevaluation handelt. Die SEVAL-Standards sind nicht für die Personalevaluation konzipiert.»

[9] Eine zentrale Eigenschaft der SEVAL-Standards ist gleichzeitig eine grosse Herausforderung: Sie sind für die in der Evaluationspraxis Tätigen nicht verbindlich, und ihre Beachtung lässt sich nur schwer durchsetzen. Dies haben die SEVAL-Standards mit denjenigen anderer Evaluationsgesellschaften durchaus gemeinsam. Für die für Deutschland und Österreich verabschiedeten Standards der DeGEval – Gesellschaft für Evaluation e. V. (Deutsche Gesellschaft für Evaluation 2002, DeGEval – Gesellschaft für Evaluation e. V. 2016) schränken BÖTTCHER und HENSE (2015, 106) deren möglichen Wirkungsgrad nämlich ein, da sie «kaum explizit regulierend wirksam werden. Da mit ihnen keinerlei Sanktionierungsmechanismen, etwa bei eklatanten Verstößen gegen die Standards guter Praxis, verbunden sind (...), bleibt die Einhaltung der Standards letztlich freiwillig».

[10] Angesichts dieser Umstände stellt sich die Frage nach der tatsächlichen Anwendung dieser Regeln guter Praxis. Dazu gibt es einige Hinweise. So findet sich mittlerweile in vielen Ausschreibungstexten für Evaluationsprojekte, dass die Einhaltung der Evaluationsstandards gefordert wird. Entsprechend wird in Offerten deren Einhaltung auch versprochen, allerdings wird nur in seltenen Fällen ausgeführt, was genau darunter zu verstehen ist. Ähnlich verhält es sich mit dem auf den Homepages vieler Evaluationsinstitutionen und -büros zu findenden Versprechen, die eigene Evaluationsarbeit an den Evaluationsstandards zu orientieren. Aber auch hier wird selten ersichtlich, wie genau sich das in der Evaluationspraxis zeigt. Entsprechend stellt WIDMER (2011, 27f) fest, dass in der Evaluationsarbeit oft frei von Konsequenzen auf die Standards Bezug genommen wird. Ausführliche, wenngleich exemplarische, auf einzelne Projekte fokussierte Darstellungen zur Nutzung der Standards findet man immer wieder (z. B. RIEDER 2009), auch zum Zweck der nachträglichen Qualitätsüberprüfung im Sinne einer Meta-Evaluation. Weniger bekannt bleibt hingegen, wie systematisch die Standards in der Praxis tatsächlich Anwendung finden.

[11] Für die frühen Jahre der SEVAL-Standards stellte LÄUBLI LOUD (2004) deren Nutzung vor allem bei der Vertragsgestaltung, bei Meta-Evaluationen und im Kontext der Lehre fest. Sie kam aber ebenso zu dem Schluss, dass der Bekanntheitsgrad und die Nutzung auf eine kleine Anzahl einzelner Personen primär innerhalb der SEVAL beschränkt waren. Knapp zehn Jahre später berichtete BALZER (2013) von einer Mitgliederbefragung bei der SEVAL, die den Evaluationsstandards einen Bekanntheitsgrad bei ihren Mitgliedern von nahezu 100 Prozent bescheinigte. Knapp 50 Prozent würden die Standards ab und zu und knapp 40 Prozent sogar regelmässig nutzen. Die Nutzungsarten sind dabei sehr vielfältig. Ein Hauptnutzen würde aber darin bestehen, «dass die Standards einen allgemeinen Orientierungsrahmen liefern und als Referenz für die Qualität von Evaluationen dienen können». So liefern die Standards einen «Gedankenanstoss», bilden eine «mentale Richtlinie» oder «man behält sie im Hinterkopf» (BALZER 2013, 452). Dies spricht für eine mehr implizite als explizite Nutzung der Standards.

[12] Noch schwieriger zu beantworten ist die Frage, wie es mit der Nutzung der Standards durch Personen ausserhalb der Fachgesellschaften aussieht. Es ist plausibel anzunehmen, dass die

SEVAL-Standards ausserhalb der SEVAL weniger bekannt sein und genutzt werden dürften. Dennoch gibt es auch ausserhalb der SEVAL eine gelebte Evaluationspraxis und Qualitätsvorstellungen, die z. B. aus anderen Bezugsdisziplinen stammen. Inwiefern sich diese von den in den Evaluationsstandards formulierten Vorstellungen unterscheiden, ist allerdings nicht klar. Aussagen zur guten Praxis von Evaluation sind somit nicht zwingend explizit an die Evaluationsstandards gebunden, ihnen eventuell aber nicht unähnlich. Es ist zudem denkbar, dass sich Evaluierende in ihrer Praxis an das in den Evaluationsstandards Formulierte halten, ohne sich dessen explizit bewusst zu sein.

[13] Für ein Verständnis davon, wie in Evaluationsstandards formulierte Grundsätze umgesetzt werden, kommt somit der Analyse der konkreten Evaluationspraxis Bedeutung zu, die sowohl eine implizite als auch eine explizite Nutzung der Standards ebenso einschliesst wie mögliche andere Regeln guter Praxis.

3. Studien zur Evaluationspraxis

3.1. Forschungsfragen

[14] Die folgenden beiden Forschungsfragen sollen die Evaluationspraxis entsprechend beleuchten:

- a. Welche wichtigen Bedingungen erfolgreicher Evaluationsprojekte können durch in der Praxis mit Evaluationsprojekten befasste Personen identifiziert werden?
- b. Inwieweit verändert sich diese Einschätzung über die Jahre?

[15] Je nachdem, wie sich die Evaluationspraxis darstellt, könnte das wiederum Einfluss auf die Gestaltung von Evaluationsstandards haben. Entsprechend widmet sich die dritte Forschungsfrage dem Verhältnis von Evaluationspraxis und Evaluationsstandards:

- c. Welchen Bezug haben Bedingungen erfolgreicher Evaluationsprojekte zu den Evaluationsstandards?

[16] Mit Hilfe der vierten Forschungsfrage sollen mögliche Defizite der konkreten Evaluationspraxis identifiziert werden, unabhängig davon, ob sie in den Evaluationsstandards thematisiert sind:

- d. Inwieweit besteht bezüglich dieser Bedingungen Handlungsbedarf in der Evaluationspraxis?

3.2. Durchführung

[17] Zur Beantwortung der vier Forschungsfragen werden Daten von zwei aufeinander bezogenen deutschsprachigen Studien verwendet, die Bedingungen erfolgreicher Evaluationsprojekte empirisch untersuchen.

[18] In der ersten, zwischen 2002 und 2003 durchgeführten Delphie-Studie (vgl. ausführlich in BALZER 2005) wurde in einer ersten, qualitativ orientierten Befragungsrunde untersucht, welche Bedingungen erfolgreicher Evaluationsprojekte Evaluationsexpertinnen und -experten nennen, wenn man ihnen eine Frage danach offen stellt und ihnen keinerlei Antwortvorgaben macht. In einer zweiten, quantitativ orientierten Befragungsrunde wurden sowohl die Wichtigkeit als auch die aktuelle Praxis der in der ersten Runde genannten Bedingungen näher untersucht. In der zweiten, im Jahr 2017 durchgeführten Studie wurde die quantitativ orientierte Befragungsrunde der ersten Studie wiederholt, d. h. exakt die gleichen, 2005 ermittelten Erfolgsbedingungen wurden Evaluationsexpertinnen und -experten erneut zur Einschätzung vorgelegt.

[19] Das Hauptinteresse galt den folgenden zwei Fragen, die zu jeder der insgesamt 103 vorgestellten Bedingungen erfolgreicher Evaluationsprojekte gestellt wurden:

«Wie bewerten Sie diese Bedingung?»

1 = ohne diese Bedingung ist kein erfolgreiches Evaluationsprojekt denkbar

2 = sehr wichtige Bedingung

3 = wichtige Bedingung

4 = relevante Bedingung, ohne die man aber auskommen kann

5 = unbedeutende Bedingung

6 = diese Bedingung ist sogar kontraproduktiv

«Wie sieht Ihrer Erfahrung nach die aktuelle Evaluationspraxis für diese Bedingung aus?»

I = wird zu viel beachtet

II = wird im richtigen Ausmass betrachtet

III = wird zu wenig beachtet

[20] Für kompetente Auskünfte zur Evaluationspraxis waren Einschätzungen von Personen gefragt, die als Expertinnen und Experten für Evaluation gelten können. Der Begriff Expertin/Experte wird dabei bewusst breit ausgelegt: Als solche werden in diesem Kontext alle Personen angesehen, die konkrete Erfahrungen mit Evaluationsprojekten gesammelt haben. Sie sind damit entweder evaluationsthematisch bewandert oder zumindest Expertinnen und Experten für ihre eigenen Evaluationserfahrungen.

[21] Angesprochen waren deswegen:

- Personen, die Evaluationsprojekte selbst durchführen, als Expertinnen und Experten für die Verknüpfung von Theorie und Praxis im Rahmen des jeweiligen Projektkontextes
- Personen, die Evaluationsprojekte in Auftrag geben, als Expertinnen und Experten für die Anforderungen, die an professionelle Evaluationsprojekte gestellt werden, damit sie finanziert werden
- Betroffene von Evaluationsprojekten als Expertinnen und Experten für ihre Bedürfnisse und Wünsche insbesondere bezüglich der Ergebnisverwendung
- Befragte in Evaluationsprojekten als Expertinnen und Experten für ihre Bedürfnisse und Wünsche bezüglich Evaluationsinhalten und -methoden

- Personen im Bereich der Evaluationsforschung (theoretisch wie praktisch) ohne direkten Projektbezug für die theoretische Beschreibung des Feldes Evaluation

[22] Die Bestimmung einer Grundgesamtheit für eine auf diese Art definierte Expertengruppe ist nicht möglich, da die Personenanzahl, deren Merkmale und weitere Kennwerte nicht einmal annähernd bekannt sind. Dementsprechend ist keine Ziehung einer repräsentativen Stichprobe möglich. Es musste aber das Ziel sein, eine «kritische» Masse an Expertise zu erreichen, die es erlaubt, eine möglichst grosse Bandbreite von Expertenmeinungen und eine mögliche Heterogenität abzubilden. Der Rekrutierungserfolg ist ausführlich zu beschreiben, damit nachvollziehbar wird, auf welcher Datenbasis die resultierenden Aussagen beruhen.

[23] Um solchermassen definierte Expertinnen und Experten zu finden, wurde in beiden Studien sehr vergleichbar vorgegangen und konkret Folgendes unternommen:

- Bekannte Expertinnen und Experten wurden per E-Mail oder Brief angeschrieben und um eine Studienteilnahme gebeten.
- Informationen über die Studie (u. a. ein Link auf eine Projektinternetseite) wurden über evaluationsbezogene Mailinglisten und Newsgroups und in der zweiten Studie auch über relevante Social-Media-Kanäle verbreitet.
- Informationen über die Studie wurden bei nationalen und internationalen Evaluationsgesellschaften und anderen, mit dem Thema Evaluation befassten Organisationen und Vereinen platziert.
- Gefundene und zur Teilnahme bereite Expertinnen und Experten wurden nach dem Schneeballsystem um Namen und Kontaktinformationen möglicher weiterer Interessentinnen und Interessenten gebeten (ähnlich dem Prinzip der «co-nomination» in Nedeva et al. 1996).
- Über (Internet-)Recherchen wurden weitere Personen gesucht und ebenfalls angeschrieben.

[24] Die Teilnahme an der Studie 2003 konnte auf verschiedene Arten erfolgen: Zum einen wurde sie mit Hilfe von Online-Fragebogen ermöglicht. Da der Zugang zum Internet zu Studienzwecken damals noch nicht so verbreitet war wie heute und um niemanden wegen technischer Barrieren auszuschliessen, gab es darüber hinaus eine computergestützte Offline-Variante sowie Papierfragebogen zum Download und zum postalischen Versand. 2017 wurde die Studie ausschliesslich als Online-Fragebogen durchgeführt.

3.3. Realisierte Stichproben

[25] Angaben von 455 Personen aus der ersten bzw. von 242 Personen aus der zweiten Studie bilden die empirische Grundlage. Um den Ursprung der Informationen nachvollziehen zu können, werden in der nachfolgenden Tabelle Stichprobeninformationen zum Geschlecht, zum Alter, zur Herkunft, zum höchsten Bildungsabschluss bzw. Titel, zum wichtigsten Evaluationsthema sowie zur inhaltlichen Verbindung zur Evaluation in den letzten 12 Monaten angegeben.

	Studie 1 (N = 455)	Studie 2 (N = 242)
Geschlecht	164 (36,4%) weiblich	128 (53,8%) weiblich
	286 (63,6%) männlich	110 (46,2%) männlich
Alter	5 ohne Angabe	4 ohne Angabe
	17–69 Jahre	25–73 Jahre
Herkunft	M = 41,1 (SD = 10,9)	M = 48,5 (SD = 10,7)
	339 (75,3%) Deutschland	111 (45,9%) Deutschland.
	62 (13,8%) Schweiz	109 (45,0%) Schweiz
	38 (8,4%) Österreich	15 (6,2%) Österreich
	11 (2,4%) sonstig	7 (2,9%) sonstig
	5 ohne Angabe	
Höchster Bildungsabschluss bzw. Titel	211 (46,4%) akademisch	138 (57,3%) akademisch
	140 (30,8%) Doktorat	74 (30,7%) Doktorat
	67 (14,7%) Professur	26 (10,8%) Professur
	37 (8,1%) sonstig	3 (1,2%) sonstig
		1 ohne Angabe
Wichtigstes Evaluationsthema	161 (49,2%) Bildung	110 (46,4%) Bildung
	42 (12,8%) Methoden der Evaluation	26 (11,0%) internationale Entwicklungszusammenarbeit
	35 (10,7%) Medizin/Gesundheitswesen/Pflege	23 (9,7%) Medizin/Gesundheitswesen/Pflege
	19 (5,8%) Politik	19 (8,0%) Politik
	15 (4,6%) Wirtschaft	16 (6,8%) Soziale Arbeit
	14 (4,3%) EDV/Telekommunikation/Internet	10 (4,2%) Wirtschaft/Finanzwesen/Ökonomie
	2 (0,6%) allgemeine Evaluationstheorie	7 (3,0%) Natur/Umwelt/Energie
	2 (0,6%) Handwerk/Technik	3 (1,3%) Justiz
	1 (0,3%) Justiz	1 (0,4%) Handel/Technik
	1 (0,3%) Militär	22 (9,3%) sonstig
	35 (10,7%) sonstig	5 ohne Angabe
	128 ohne Angabe	
Inhaltliche Verbindung zur Evaluation in den letzten 12 Monaten	263 (57,8%) Evaluationen selbst durchgeführt	153 (63,2%) Evaluationen selbst durchgeführt
	70 (15,4%) Evaluationen in Auftrag gegeben	53 (21,9%) Evaluationen in Auftrag gegeben
	188 (41,3%) in Evaluationen befragt	74 (30,6%) in Evaluationen befragt
	296 (65,1%) Evaluationsergebnisse genutzt	166 (68,6%) Evaluationsergebnisse genutzt
	206 (45,3%) beratend in Evaluationen tätig	126 (52,1%) beratend in Evaluationen tätig
	127 (27,9%) im Bereich der Evaluation ausgebildet	77 (31,8%) im Bereich der Evaluation ausgebildet
	37 (8,1%) sonstige Verbindung zur Evaluation	14 (5,8%) sonstige Verbindung zur Evaluation

Tabelle 1: realisierte Stichproben

3.4. Ergebnisse: Erfolgsbedingungen gestern und heute

[26] Mit der ersten Analyse zur Beantwortung der Forschungsfragen a und b wird bestimmt, wie die befragten Expertinnen und Experten die Wichtigkeit der 103 ermittelten Erfolgsbedingungen einschätzen. Dies erfolgt über den Mittelwert aller entsprechenden Bewertungen (von «1 = ohne diese Bedingung ist kein erfolgreiches Evaluationsprojekt denkbar» bis «6 = diese Bedingung ist sogar kontraproduktiv», vgl. Ziff. 3.1). Die Datenbasis für die folgenden Darstellungen bilden Angaben von 455 Personen für 2003 und 242 Personen für 2017. Dabei schwankt diese Anzahl von Bedingung zu Bedingung geringfügig aufgrund vereinzelter fehlender Werte.

[27] Es können 33 von 103 Erfolgsbedingungen identifiziert werden, die entweder in der Studie von 2003 oder von 2017 einen Mittelwert von 2,0 aufweisen. Dieser Mittelwert steht für eine als mindestens sehr wichtig eingeschätzte Bedingung. Die nachfolgende Tabelle gibt die Mittelwerte (M) dieser 33 Erfolgsbedingungen in der Reihenfolge ihrer Wichtigkeit aus der 2017er-Befragung sowie deren Standardabweichungen (SD) an: je niedriger der Mittelwert, desto höher die Wichtigkeit der genannten Bedingung. Die Ziffer vor jeder Erfolgsbedingung repräsentiert die im Fragebogen verwendete Nummerierung, die im weiteren Text verwendet wird.

Erfolgsbedingung	2003		2017	
	M	SD	M	SD
2 – klare Definition des Evaluationsgegenstandes	1,31	0,65	1,36	0,68
36 – Unbestechlichkeit des Evaluators	1,40	0,76	1,37	0,75
8 – klare, realistische Evaluationsziele und -fragestellungen	1,50	0,73	1,40	0,69
76 – korrekte und methodisch saubere Datenauswertung	1,46	0,68	1,55	0,79
37 – Akzeptanz und Glaubwürdigkeit des Evaluators	1,62	0,76	1,58	0,64
85 – der Evaluationsbericht beschreibt transparent den Evaluationsgegenstand und seinen Kontext sowie die Evaluationsziele, -verfahren und -ergebnisse	1,71	0,76	1,60	0,75
65 – korrekte und methodisch saubere Evaluationsdurchführung	1,59	0,79	1,63	0,75
63 – Beachtung rechtlicher Grundsätze	1,73	0,91	1,63	0,82
62 – Beachtung ethischer Grundsätze	1,72	0,94	1,65	0,80
39 – Objektivität und Neutralität des Evaluators	1,83	0,98	1,65	0,95
40 – Unabhängigkeit des Evaluators	1,77	0,97	1,66	0,88
59 – eindeutiges, angemessenes Evaluationsdesign	1,70	0,77	1,68	0,81
29 – hohe Evaluations-Expertise des Evaluators	1,81	0,82	1,69	0,74
80 – die Ergebnisberichterstattung erfolgt unparteiisch	1,77	0,87	1,69	0,80
82 – sensibler, vertraulicher, sachlicher Umgang mit Ergebnissen	1,73	0,87	1,72	0,83
21 – Offenheit des Auftraggebers gegenüber der Evaluation und ihren Ergebnissen	1,85	0,86	1,72	0,75
81 – nachvollziehbare, adressatengerechte Ergebnisdarlegung	1,75	0,75	1,72	0,74

91 – Ergebnisinterpretationen werden ausdrücklich begründet	1,81	0,91	1,74	0,81
93 – der Ergebnisbericht beinhaltet keine vorschnellen Schlussfolgerungen	1,68	0,82	1,74	0,84
92 – im Ergebnisbericht werden Befunde, Interpretationen und Empfehlungen getrennt	1,60	0,89	1,74	0,93
6 – klare Definition der Evaluationsobjekte	1,54	0,73	1,75	0,88
61 – Beachtung des Datenschutzes	1,77	1,04	1,76	0,94
12 – schriftlich fixierte Evaluationsziele und -fragestellungen	1,90	0,89	1,82	0,89
72 – Beachtung einschlägiger Qualitätsstandards	1,85	0,82	1,85	0,78
102 – Ergebnisse werden diskutiert	1,98	0,99	1,88	0,82
14 – klare Bewertungsgrundlagen und -indikatoren	1,70	0,81	1,88	0,95
3 – klare Definition des Kontextes, in dem der Evaluationsgegenstand angesiedelt ist	1,76	0,79	1,91	0,80
78 – (potentiell) nützlich-konstruktiv Ergebnisse	2,12	1,03	1,94	0,89
79 – Ergebnisse sind vollständig und fair	2,01	0,83	1,94	0,87
71 – Transparenz über den Evaluationsprozess	1,96	0,89	1,96	0,84
32 – Selbstreflexionsfähigkeit des Evaluators	2,09	0,88	1,98	0,83
53 – Vorhandensein ausreichender Ressourcen	2,04	0,90	1,99	0,75
86 – Transparenz über die Verwendung der Ergebnisse wird hergestellt	1,97	0,85	2,13	0,92

Tabelle 2: Wichtigkeit der Erfolgsbedingungen 2003 und 2017

[28] Im Vergleich der Ergebnisse von 2003 und 2017 lässt sich festhalten, dass sich die Einschätzungen in 14 Jahren kaum verändert haben. Die nachfolgende Zusammenfassung gilt somit für beide Erhebungszeiträume in gleicher Weise. Die Einschätzungen der Befragten zu Bedingungen, die für den Erfolg von Evaluationsprojekten besonders relevant sind, lassen sich thematisch gruppiert wie folgt zusammenfassen (in Klammern die in der Tabelle verwendete Nummer):

[29] Um ein erfolgreiches Evaluationsprojekt realisieren zu können ist es besonders wichtig zu klären, was (2), mit welchem Ziel und mit welchen Fragestellungen (8) – und dies schriftlich fixiert (12) – bei wem (6), auf welcher Bewertungsgrundlage (14) und in welchem Kontext (3) evaluiert wird.

[30] Der Auftraggeber muss offen sein gegenüber der Evaluation und den resultierenden Ergebnissen (21).

[31] Selbstverständlich spielen die Evaluierenden für das Gelingen einer Evaluation eine besondere Rolle. Besonders Unbestechlichkeit (36), Akzeptanz und Glaubwürdigkeit (37), Unabhängigkeit (40), hohe Evaluationsexpertise (29), Objektivität und Neutralität (39) sowie Selbstreflexionsfähigkeit (32) zeichnen erfolgreich Evaluierende aus.

[32] Darüber hinaus ist es notwendig, dass eine Evaluation professionell abläuft. Dazu gehören eine korrekte und methodisch saubere Evaluationsdurchführung (65) und ebensolche Datenauswertung (76), ein eindeutiges, angemessenes Evaluationsdesign (59), die Beachtung ethischer (62), rechtlicher (63) und datenschutzbezogener Grundsätze (61), die Beachtung einschlägiger Qualitätsstandards (72), die Transparenz über den Evaluationsprozess (71) sowie das Vorhandensein ausreichender Ressourcen (53).

[33] Die Ergebnisse einer Evaluation und deren Berichterstattung können zum Erfolg besonders dann beitragen, wenn im Ergebnisbericht Befunde, Interpretationen und Empfehlungen getrennt werden (92), wenn der Ergebnisbericht keine vorschnellen Schlussfolgerungen beinhaltet (93), wenn der Ergebnisbericht den Evaluationsgegenstand und seinen Kontext sowie die Evaluationsziele, -verfahren und -ergebnisse transparent beschreibt (85), wenn mit Ergebnissen sensibel und vertraulich umgegangen wird (82), wenn die Ergebnisdarlegung nachvollziehbar und adressatengerecht stattfindet (81), wenn die Ergebnisberichterstattung unparteiisch erfolgt (80), wenn Ergebnisinterpretationen ausdrücklich begründet werden (91), wenn Transparenz über die Verwendung der Ergebnisse hergestellt wird (86), wenn Ergebnisse diskutiert werden (102), wenn (potenziell) nützlich-konstruktive Ergebnisse vorhanden sind (78) und wenn Ergebnisse vollständig und fair sind (79).

3.5. Ergebnisse: Erfolgsbedingungen und Evaluationsstandards

[34] Die bisher dargestellten Ergebnisse zeigen u. a., dass es damals wie heute viele Erfolgsbedingungen gibt, die die Befragten als besonders wichtig für ein erfolgreiches Evaluationsprojekt einschätzen.

[35] Nun stellt sich mit Forschungsfrage c die Frage, inwiefern sich die als besonders wichtig eingeschätzten Erfolgsbedingungen in den Evaluationsstandards als Korpus guter, professioneller Evaluationspraxis wiederfinden. Ein systematischer Vergleich der empirisch bestimmten Erfolgsbedingungen mit den in drei Gruppen (A – Allgemeine Grundprinzipien, B – Planung und Durchführung, C – Bewertung und Ergebnisvermittlung) gegliederten einzelnen SEVAL-Standards (SEVAL – Schweizerische Evaluationsgesellschaft 2016) zeigt dabei viele Gemeinsamkeiten.

[36] Es wurde beschrieben, dass es für ein erfolgreiches Evaluationsprojekt als Grundlage und Voraussetzung besonders wichtig ist zu klären, was (2), mit welchem Ziel und mit welchen Fragestellungen (8) – und dies schriftlich fixiert (12) – bei wem (6), auf welcher Bewertungsgrundlage (14) und in welchem Kontext (3) evaluiert wird. Diese Sachverhalte finden sich alle in den Standards wieder, namentlich in B4, B1, B1/B4, A3, C2 sowie B2.

[37] Weiter muss der Auftraggeber offen sein gegenüber der Evaluation und den resultierenden Ergebnissen (21). Dieser Sachverhalt findet sich bei den Standards in A1 (beschränkt sich dort aber nicht nur auf die Auftraggeber).

[38] Die Evaluierenden spielen für das Gelingen einer Evaluation eine besondere Rolle. Während deren hohe Evaluationsexpertise (29) in Standard A6, deren Unabhängigkeit (40) in A4, deren Unbestechlichkeit (36) in A12 sowie deren Objektivität und Neutralität (39) in A11 benannt ist (wobei die Standards diese Eigenschaften nicht nur auf Evaluierende, sondern auf alle Beteiligten und Betroffenen beziehen), finden die Akzeptanz und Glaubwürdigkeit (37) sowie die Selbstreflexionsfähigkeit (32) der Evaluierenden zunächst kein explizites Pendant in den ausformulierten Standards. Allerdings wird man einerseits in den Erläuterungen zu den Standards für Evaluation fündig (SEVAL – Schweizerische Evaluationsgesellschaft 2017a). Die Erläuterungen zu Standard A6 benennen wichtige Eigenschaften von Evaluierenden (und anderen am Evaluationsprozess Beteiligten) wie z. B. soziale Kompetenzen. Andererseits hat die SEVAL der Relevanz des Themas entsprechend erforderliche Kompetenzen von Evaluierenden in einem separaten Text ausführlich

beschrieben (vgl. ausführlich SEVAL – Schweizerische Evaluationsgesellschaft 2012). Die dort beschriebene Liste ist noch erheblich detaillierter.

[39] Eine weitere wichtige Erfolgsbedingung einer Evaluation ist, dass sie professionell abläuft. Dazu gehört eine korrekte und methodisch saubere Evaluationsdurchführung (65), was sich im Standard B6 verorten lässt. Eine korrekte und methodisch saubere Datenauswertung (76) wird in Standard B8 (der die Datenerhebung noch explizit einschliesst) angesprochen, ein eindeutiges, angemessenes Evaluationsdesign (59) in Standard B5, die Beachtung ethischer Grundsätze (62) in Standard A10, die Beachtung rechtlicher Grundsätze (63) in Standard A8, der Datenschutz (61) in Standard A9 und die Transparenz über den Evaluationsprozess (71) in Standard A2. Das Vorhandensein ausreichender Ressourcen (53) findet in den Standards keine Entsprechung, ebenso wenig die Beachtung einschlägiger Qualitätsstandards (72). Es wäre in Bezug auf Letzteres aber auch etwas merkwürdig, wenn in den Standards z. B. deren eigene Einhaltung als expliziter Standard formuliert wäre.

[40] Besonders oft wurden Elemente der Berichterstattung als wichtige Erfolgsbedingungen thematisiert. Dieses Thema ist bei den Erfolgsbedingungen breiter abgedeckt als in den Standards. Eine Entsprechung im Standard A2 findet die Erfolgsbedingung, im Ergebnisbericht den Evaluationsgegenstand und seinen Kontext sowie die Evaluationsziele, -verfahren und -ergebnisse transparent zu beschreiben (85). In A9 ist (allerdings nicht nur auf Ergebnisse beschränkt) angesprochen, dass mit Ergebnissen sensibel und vertraulich umgegangen wird (82), in C4, dass die Ergebnisdarlegung nachvollziehbar und adressatengerecht stattfindet (81), ebenso in C4, dass die Ergebnisberichterstattung unparteiisch erfolgt (80), in C2, dass Ergebnisinterpretationen ausdrücklich begründet werden (91), nochmals in C2, dass der Ergebnisbericht keine vorschnellen Schlussfolgerungen beinhaltet (93) sowie in C1, dass Ergebnisse vollständig und fair sind (79). Eine Entsprechung findet sich so ungefähr in C3, dass (potenziell) nützlich-konstruktive Ergebnisse vorhanden sind (78) – im Standard geht es allerdings eher in Richtung von zu formulierenden Empfehlungen. Eine Entsprechung findet sich darüber hinaus annähernd in C6, dass Transparenz über die Verwendung der Ergebnisse hergestellt wird (86). Allerdings geht die Anforderung dieser Erfolgsbedingung insofern über die des Standards C6 hinaus, dass im Standard die Transparenz nur betreffend die Ergebnisse und nur gegenüber den Beteiligten und Betroffenen herzustellen ist, wohingegen in der Erfolgsbedingung nicht nur die Ergebnisse selbst offenzulegen sind, sondern sogar deren Verwendung. Hingegen wird in den Standards nicht explizit erwähnt, dass Ergebnisse einer Evaluation und deren Berichterstattung zum Erfolg besonders dann beitragen können, wenn im Ergebnisbericht Befunde, Interpretationen und Empfehlungen getrennt (92) und wenn Ergebnisse diskutiert werden (102). Insgesamt haben die Standards bezüglich der Ergebnisse eine andere Flughöhe. Ähnlich wie bei den Eigenschaften der Evaluierenden sind die Standards nicht geeignet – und haben auch gar nicht den Anspruch –, um das Thema der Berichterstattung in diesem Detailliertheitsgrad abzubilden. Die Bedeutung dieses Aspektes für den Erfolg eines Evaluationsprojektes wird aber dadurch ersichtlich, dass die Standardgruppe C der Bewertung und Ergebnisvermittlung gewidmet ist. Die revidierten Standards fassen die in der ersten Version auf verschiedene Standardgruppen verteilten Merkmale der Berichterstattung zusammen. Was hierbei im Detail zu beachten ist, wird an anderer Stelle ausführlich beschrieben und zusammengefasst (z. B. BALZER/BEYWL 2018, 129 ff).

3.6. Ergebnisse: Handlungsbedarf in der Evaluationspraxis

[41] Forschungsfrage d untersucht, inwieweit die Bedingungen erfolgreicher Evaluationsprojekte in der Einschätzung von Expertinnen und Experten in der Evaluationspraxis Beachtung finden. Wie erwähnt, wurde in beiden Studien dafür für jede der 103 Erfolgsbedingungen die Frage gestellt: «Wie sieht Ihrer Erfahrung nach die aktuelle Evaluationspraxis für diese Bedingung aus?». Es standen die Antwortalternativen «I = wird zu viel beachtet», «II = wird im richtigen Ausmass beachtet» sowie «III = wird zu wenig beachtet» zur Verfügung.

[42] Handlungsbedarf wird an dieser Stelle bescheinigt, wenn mindestens 50 Prozent der Antwortenden eine zu geringe Beachtung einer Erfolgsbedingung in der Praxis nennen (eine zu starke Beachtung wurde keiner der genannten Bedingungen in nennenswertem Umfang bescheinigt). Besonderes Gewicht erhält dieser Handlungsbedarf zudem insbesondere bei als besonders wichtig eingeschätzten Erfolgsbedingungen. Die folgende Tabelle 3 zeigt die entsprechenden Ergebnisse in absteigender Reihenfolge des Handlungsbedarfes aus der Studie 2017. Dargestellt werden solche Erfolgsbedingungen, die zuvor als besonders wichtig identifiziert worden sind, und aus dieser Auswahl nur diejenigen, die in beiden Studien einen Handlungsbedarf von mindestens 50 Prozent aufweisen.

Handlungsbedarf bei besonders wichtigen Erfolgsbedingungen	2003	2017
	in %	in %
8 – klare, realistische Evaluationsziele und -fragestellungen	68.3	62.7
2 – klare Definition des Evaluationsgegenstandes	65.3	57.3
14 – klare Bewertungsgrundlagen und -indikatoren	57.6	54.6
3 – klare Definition des Kontextes, in dem der Evaluationsgegenstand angesiedelt ist	62.2	53.2
53 – Vorhandensein ausreichender Ressourcen	59.2	51.4
86 – Transparenz über die Verwendung der Ergebnisse wird hergestellt	57.6	51.0

Tabelle 3: Handlungsbedarf bei den wichtigsten Erfolgsbedingungen

[43] Der Handlungsbedarf ist 2017 gegenüber 2003 für alle dargestellten Erfolgsbedingungen zwar gesunken, aber in den Augen der Befragten gibt es nach wie vor Verbesserungspotenzial für die aktuelle Evaluationspraxis.

[44] Für sechs als sehr wichtig identifizierte Erfolgsbedingungen besteht gemäss beiden Studien Handlungsbedarf für die Evaluationspraxis; ihnen müsste mehr Beachtung geschenkt werden. Es ist dabei sehr auffällig, dass es sich um sehr grundlegende Klärungsschritte handelt: Klärung von Evaluationsgegenstand und dessen Kontext, von Evaluationszielen und -fragestellungen sowie der Bewertungsgrundlagen und -indikatoren. Auch Ressourcen zur Durchführung von Evaluationen bleiben eine Herausforderung sowie der Wunsch, Transparenz über die Verwendung der Ergebnisse herzustellen.

[45] Die beiden Erfolgsbedingungen «klare Definition des Evaluationsgegenstandes» (vergleichbar mit Standard B4) sowie «klare und realistische Evaluationsziele und Fragestellungen» (vergleichbar mit Standard B1) gehören darüber hinaus zu den drei wichtigsten Erfolgsbedingungen überhaupt. Mit anderen Worten wird Handlungsbedarf an besonders zentralen Schritten im Evaluationsprozess identifiziert.

3.7. Methodische Anmerkungen

[46] Einschränkend bei der Betrachtung der Ergebnisse könnte sein, dass beide Studien keine Repräsentativität ihrer Stichproben beanspruchen können. Da die Bestimmung der Grundgesamtheit aller Evaluationsexpertinnen und -experten im in den Studien verwendeten Sinn nicht möglich ist und somit weder eine Zufallsauswahl noch eine Quotierung oder Gewichtung an bestimmten Merkmalen vorgenommen werden kann, sind allgemeingültige Aussagen über die gesamte Evaluations-Community nicht möglich. Auch liegen ausschliesslich deutschsprachige Daten vor. Mögliche Einschränkungen bezüglich der vorliegenden Stichproben sind auch für die Analyse von veränderten Einschätzungen der Erfolgsbedingungen über die Zeit relevant: Da sich einige Merkmale in beiden Stichproben unterscheiden, es gleichzeitig aber im Dunkeln bleibt, ob sich hier eine veränderte Evaluations-Community 2003 versus 2017 zeigt oder ein unterschiedlicher Erfolg bei der Teilnehmendenrekrutierung, stehen solche Vergleiche unter einem bestimmten Vorbehalt. Dennoch gehen die Analysen weit über das Stadium der sonst nur möglichen Spekulation hinaus. Detaillierte Informationen zu den Befragten machen in jedem Fall nachvollziehbar, auf wessen Aussagen alle Analysen beruhen.

[47] Darüber hinaus betrachten beide Studien die Evaluationspraxis global und übergreifend. Ein solch übergreifender Blick über viele Evaluationssituationen hinweg hat Vorteile, allerdings bringt er auch Probleme mit sich: Evaluationen finden immer in bestimmten Situationen mit ihnen eigenen Rahmenbedingungen statt, die es nicht immer erlauben, Bedingungen «über einen Kamm zu scheren». So kann es durchaus sein, dass die eine oder andere Bedingung im konkreten Handlungskontext ganz anders zu bewerten wäre.

[48] Ein weiteres Problem betrifft die verwendete Fachsprache. Die Evaluationsterminologie und die Begriffsverwendung verändern sich über die Zeit. Manche Begriffe der Studie 2003 klingen heute etwas altmodisch oder werden in der Zwischenzeit präziser verwendet. So hat es sich in vielen Evaluationsbereichen z. B. durchgesetzt, von Evaluationszwecken (statt-zielen) zu sprechen. Auch der Begriff Evaluationsobjekt ist heute eher unüblich. Darüber hinaus ist eine gendersensitive Sprache heute selbstverständlicher als noch 2003. Der Entscheid, die Erfolgsbedingungen nicht an neuere Sprachentwicklungen anzupassen war schlussendlich eine Abwägung zwischen Vergleichbarkeit und Begriffsaktualität – um Ersteres nicht unmöglich zu machen, wurden sprachliche Ungenauigkeiten in Kauf genommen.

[49] Trotz der genannten Einschränkungen erlauben es die beschriebenen Daten, einen etwas anderen, auf Empirie gestützten Blick auf die Evaluationspraxis zu werfen.

4. Einordnung in die Diskussion zur Professionalisierung der Evaluation in der Schweiz

[50] In diesem abschliessenden Kapitel wird versucht, die Ergebnisse in die aktuell geführte Diskussion zur Professionalisierung der Evaluation in der Schweiz einzuordnen.

[51] Bemühungen rund um eine Professionalisierung von Evaluation sind weltweit festzustellen (vgl. MEYER 2015). Für die Evaluation in der Schweiz entwarf der damalige SEVAL-Präsident eine Roadmap für deren Professionalisierung (SANGRA 2012) und stellte fest, dass schon einiges erreicht wurde: So ist seit den 1970er-Jahren eine Erwerbstätigkeit für Evaluierende gestützt auf öffentliche Aufträge zu verzeichnen. Es existieren Angebote zur Aus- und Weiterbildung in Evaluation. Insbesondere an der Universität Bern werden etablierte Aufbaustudiengänge zur

Grundlegung einer spezifischen Wissensbasis für Evaluation angeboten. Deren Beitrag zur Professionalisierung von Evaluation thematisieren BEYWL, BESTVATER und MÜLLER (2011). Diverse Institutionen bieten Weiterbildungen an, und in verschiedenen grundständigen Studiengängen (z. B. Psychologie) gibt es Vertiefungen zu Evaluation. Man findet in der Schweiz diverse Forschungsaktivitäten über Evaluation, zuletzt zusammengefasst in WIDMER et al (2016). 1996 wurde die Schweizerische Evaluationsgesellschaft SEVAL gegründet, in der sich in unterschiedlichen Rollen mit Evaluation befassende Personen und Institutionen politikfeldübergreifend austauschen. Über die eingangs schon erwähnte Bundesverfassung, über rechtliche Erlasse sowie verschiedene Verwaltungseinheiten ist Evaluation institutionalisiert. Ein weiteres Element von Professionalisierung stellen fachspezifische Kommunikationskanäle dar. Mit dem «forum-evaluation», das als deutschsprachige, elektronische Diskussionsliste rund um das Thema Evaluation auch in der Schweiz Verwendung findet, sowie mit wissenschaftlichen und fachlichen Evaluationszeitschriften (insbesondere «LeGes – Gesetzgebung und Evaluation» sowie «Zeitschrift für Evaluation (ZfEv)») stehen solche Kanäle zur Verfügung.

[52] Ein zentrales Element auf dem Weg zur Professionalisierung ist darüber hinaus die Definition von Regeln professioneller Evaluation, von Evaluationsstandards, die Qualitätsanforderungen an gute Evaluationen formulieren. Hier wurde die SEVAL schon vor vielen Jahren aktiv und legte 2001 die erste Version ihrer Standards (WIDMER/LANDERT/BACHMANN 2001) vor. Deren erste Revision folgte 2016 (SEVAL – Schweizerische Evaluationsgesellschaft 2016). Die empirischen Ergebnisse dieses Beitrages zeigen, dass die neuen SEVAL-Standards ziemlich gut dem entsprechen, was Evaluationsexpertinnen und -experten als Bedingungen erfolgreicher Evaluationsprojekte ansehen. Eine Ausnahme bildet mit der Forderung nach dem Vorhandensein ausreichender Ressourcen eine Erfolgsbedingung, die sicherlich in nahezu jedem Evaluationsprojekt ohnehin diskutiert wird. Und wenn es um Eigenschaften von Evaluierenden und um die Berichtslegung geht, gehen die Erfolgsbedingungen mehr ins Detail als die Standards.

[53] SANGRA (2012) skizzierte weitere mögliche Schritte zur Weiterentwicklung des Evaluationsberufs, namentlich ein professionelles SEVAL-Sekretariat, eine harmonisierte und formalisierte Umsetzung und Finanzierung von Evaluationen in der Verwaltung, eine Verbesserung der SEVAL-Standards, das Bereitstellen von Mediationsangeboten als interne Regulierungsmechanismen bei Konflikten zwischen Auftragnehmern und -gebern, eine Definition der für einen Beruf notwendigen Kompetenzen, die Definition und Anerkennung der notwendigen Aus- und Weiterbildungen und gegebenenfalls auch die Zertifizierung dieser Kompetenzen, bis hin zu einer Bundes- oder kantonalen Anerkennung der Titel und der Ausbildung von Evaluatorinnen und Evaluatoren. Gemäss RÜEFLI (2012, 414) «scheint sich die Evaluation in der Schweiz in einer interessanten Zwischenphase zu befinden, in der einige Schritte hin zu einer Profession bereits gegangen sind.» «Sie kann jedoch aktuell erst bedingt als Profession im soziologischen Sinn bezeichnet werden» (ebd., 420).

[54] Seit diesen Feststellungen ist wiederum einiges passiert. So wurden z. B. Kompetenzanforderungen an Evaluierende (SEVAL – Schweizerische Evaluationsgesellschaft 2012), ein Kompetenzrahmen für Evaluationsmanagerinnen und -manager (SEVAL – Schweizerische Evaluationsgesellschaft: Arbeitsgruppe «Kompetenzen in der Evaluation» 2014), eine Broschüre zur Beschreibung von Evaluationen und deren Nützlichkeit (SEVAL – Schweizerische Evaluationsgesellschaft 2015), sowie Empfehlungen für die Konzipierung, Etablierung, Durchführung und Verwendung von Evaluationen im Rahmen des Qualitätsmanagements von Lehrveranstaltungen im Hochschulbereich (SEVAL – Schweizerische Evaluationsgesellschaft 2018) formuliert.

[55] Mit dem Strategiepapier für 2018–2021 (SEVAL – Schweizerische Evaluationsgesellschaft 2017b) schliesst der Vorstand der SEVAL an diese Arbeiten an und will sie weiterführen. Es wird aufgezeigt, wie die Professionalisierung weiter vorangetrieben werden soll und welche Rolle die SEVAL und deren Arbeitsgruppen dabei spielen können.

[56] Aktuell werden Diskussionen um eine mögliche Beratungs- oder Vermittlungsstelle sowie um Zertifizierungen von Evaluierenden kontrovers geführt, zuletzt am SEVAL-Jahreskongress 2019, der sich unter dem Titel «Zusammen die Zukunft gestalten – Professionalisierung in der Evaluation und die Rolle der SEVAL» explizit auch diesen Themen widmete.

[57] Aus einer Mitgliederbefragung der SEVAL zum Stand der Professionalisierung der Evaluation in der Schweiz geht hervor, dass die befragten Personen diesen zwar grundsätzlich positiv einschätzen (vgl. im Detail FRIEDRICH/SCHLAUFER/MICHEL 2017). Dies betrifft sowohl die Qualität von Evaluationen und Kompetenzen von Auftragnehmenden als auch das Vorhandensein einer gemeinsamen Wissensbasis zu Evaluation innerhalb der SEVAL. Doch ein anerkannter Beruf «Evaluatorin/Evaluator» ist nach wie vor in weiter Ferne. Evaluator/Evaluatorin darf sich zurzeit jeder nennen, weder Titel noch Tätigkeit sind gesetzlich geschützt. Und genau diesen fehlenden Titelschutz erachtet eine Mehrheit der Befragungsteilnehmenden als problematisch. Eine überwiegende Mehrheit sieht denn auch weiteren Handlungsbedarf in Richtung Professionalisierung, namentlich zum Schutz von Beteiligten und Betroffenen vor schlechten Evaluationen. Um die Professionalisierung voranzutreiben, sieht ebenfalls eine deutliche Mehrheit die SEVAL in der Pflicht, ihre bisherigen Anstrengungen noch zu intensivieren.

[58] Eine weitere Herausforderung wird durch die empirischen Ergebnisse dieser Studie offenbar: Trotz der beschriebenen zahlreichen Aktivitäten in Richtung Professionalisierung sehen Evaluationsexpertinnen und -experten Handlungsbedarf bei den Ressourcen zur Durchführung von Evaluationen sowie bei der Herstellung von Transparenz über die Verwendung der Evaluationsergebnisse. Insbesondere werden aber für sehr grundlegende Klärungsschritte in Evaluationsprojekten grosse Defizite festgestellt und entsprechender Handlungsbedarf für die Praxis identifiziert: Demnach sollten der Klärung von Evaluationsgegenstand und dessen Kontext, von Evaluationszielen und -fragestellungen sowie der Bewertungsgrundlagen und -indikatoren in der Praxis mehr Aufmerksamkeit geschenkt werden, als dies aktuell geschieht. Dies ist insofern bemerkenswert, als damit absolute Grundlagen des konkreten Handelns in Evaluationsprojekten angesprochen sind. Bei der Planung nächster Schritte zur Professionalisierung von Evaluation scheint es deswegen angebracht, diese absoluten Grundlagen des Evaluierens immer im Auge zu behalten. Funktioniert die Praxis in dieser Hinsicht nicht, wird auch eine Professionalisierung nicht erfolgreich sein können.

Prof. Dr. LARS BALZER, Leiter Fachstelle Evaluation, Eidgenössisches Hochschulinstitut für Berufsbildung EHB, evaluation@lars-balzer.info.

Literaturverzeichnis

BALTHASAR, ANDREAS (2007): *Institutionelle Verankerung und Verwendung von Evaluationen: Praxis und Verwendung von Evaluationen in der schweizerischen Bundesverwaltung*. Zürich: Rüegger.

BALZER, LARS (2005): *Wie werden Evaluationsprojekte erfolgreich? – Ein integrierender theoretischer Ansatz und eine empirische Studie zum Evaluationsprozess*. Landau: Verlag Empirische Pädagogik.

- BALZER, LARS (2013): *Bekanntheitsgrad und Nutzung der Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL) sowie erste Modifikationshinweise: Ergebnisse der Mitgliederbefragung*. In: LeGes – Gesetzgebung & Evaluation 24/2. S. 439–458.
- BALZER, LARS und BEYWL, WOLFGANG (2018): *evaluiert – erweitertes Planungsbuch für Evaluationen im Bildungsbereich*. 2. Aufl. Bern: hep verlag.
- BEYWL, WOLFGANG; BESTVATER, HANNE und MÜLLER, MARIANNE (2011): *Professionalisierung der Evaluation durch Weiterbildung?* In: LeGes – Gesetzgebung & Evaluation 22/1. S. 89–109.
- BÖTTCHER, WOLFGANG und HENSE, JAN (2015): *Professionelle Evaluation oder Evaluation als Profession?* In: Hennefeld, Vera (Hrsg.): *Nachhaltige Evaluation?: Auftragsforschung zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann*. Münster: Waxmann. S. 101–119.
- BUSSMANN, WERNER (2008): *The Emergence of Evaluation in Switzerland*. In: Evaluation 14/4. S. 499–506. doi:10.1177/1356389008095491.
- BUSSMANN, WERNER (2010): *Evaluation of Legislation: Skating on Thin Ice*. In: Evaluation 16/3. S. 279–293. doi:10.1177/1356389010370252.
- BUSSMANN, WERNER (2015): *Konstanz, Wandel und Neueentwicklungen: 25 Jahre Evaluation in LeGes*. In: LeGes – Gesetzgebung & Evaluation 26/1. S. 107–125.
- DeGEval – Gesellschaft für Evaluation e. V. (2016): *Standards für Evaluation – erste Revision 2016*. Mainz.
- Deutsche Gesellschaft für Evaluation (2002): *Standards für Evaluation*. Köln: DeGEval.
- FRIEDRICH, VERENA; SCHLAUFER, CAROLINE und MICHEL, IRIS (2017): *Professionalisierung der Evaluation: Ergebnisse der SEVAL-Mitgliederbefragung*. In: LeGes – Gesetzgebung & Evaluation 28/2. S. 347–367.
- HORBER-PAPAZIAN, KATIA (Hrsg.) (2015): *Regards croisés sur l'évaluation en Suisse*. Lausanne: Presses polytechniques et universitaires romandes.
- JACOB, STEVE; SPEER, SANDRA und FURUBO, JAN ERIC (2015): *The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later*. In: Evaluation 21/1. S. 6–31. doi:10.1177/1356389014564248.
- LÄUBLI LOUD, MARLÈNE (2004): *Setting Standards and Providing Guidelines: The Means Toward What End?* In: Evaluation 10/2. S. 237–245.
- MADER, LUZIUS (2005): *Artikel 170 der Bundesverfassung: Was wurde erreicht, was ist noch zu tun?* In: LeGes – Gesetzgebung & Evaluation 16/1. S. 29–37.
- MEYER, WOLFGANG (2015): *Professionalisierung von Evaluation: ein globaler Blick*. In: Zeitschrift für Evaluation 14/2. S. 215–246.
- NEDEVA, MARIA; GEORGHIOU, LUKE; LOVERIDGE, DENIS und CAMERON, HUGH (1996): *The use of co-nomination to identify expert participants for Technology Foresight*. In: R&D Management 26/2. S. 155–168.
- RIEDER, STEFAN (2009): *Anwendung der SEVAL-Standards: Ergebnisse aus vier Fallstudien*. In: LeGes – Gesetzgebung & Evaluation 20/3. S. 387–397.
- RÜEFLI, CHRISTIAN (2012): *Evaluation und Qualitätssicherung – Rückblick auf den SEVAL-Jahreskongress 2012*. In: LeGes – Gesetzgebung & Evaluation 23/3. S. 397–422.
- RÜEFLI, CHRISTIAN (2013): *Die Revision der SEVAL-Standards – Kontext, Vorgehen und weiterführende Überlegungen*. In: LeGes – Gesetzgebung & Evaluation 24/2. S. 459–469.
- RÜEFLI, CHRISTIAN und ZWEIDLER, REINHARD (2019): *Die SEVAL-Standards und ihre Revision – ein Erfahrungsbericht*. In: Hense, Jan-Ulrich; Böttcher, Wolfgang; Kalman, Michael und Meyer, Wolfgang (Hrsg.): *Evaluation: Standards in unterschiedlichen Handlungsfeldern. Einheitliche Qualitätsansprüche trotz heterogener Praxis?* Münster: Waxmann. S. 261–276.
- Sager, Fritz; Widmer, Thomas und Balthasar, Andreas (Hrsg.) (2017): *Evaluation im politischen System der Schweiz*. Zürich: NZZ Libro.
- SANGRA, EMMANUEL (2012): *Beruf: Evaluator...* In: LeGes – Gesetzgebung & Evaluation 23/2. S. 217–221.
- SEVAL – Schweizerische Evaluationsgesellschaft: Arbeitsgruppe «Kompetenzen in der Evaluation» (2014): *Kompetenzrahmen für Evaluationsmanagerinnen und Evaluationsmanager: Aufgaben, Herausforderungen, Kompetenzen*. Bern: SEVAL. [https://www.seval.ch/app/uploads/2017/07/Brochure_SEVAL-Kompetenzen_Evaluationsmanag-d.pdf; 1.10.2019].
- SEVAL – Schweizerische Evaluationsgesellschaft (2012): *Kompetenzanforderungen in der Evaluation*. [https://www.seval.ch/app/uploads/2017/07/SEVAL_competences_d.pdf; 1.10.2019].
- SEVAL – Schweizerische Evaluationsgesellschaft (2015): *Evaluationen: Was sie sind und was sie bringen*. Fribourg. [https://www.seval.ch/app/uploads/2017/06/SEVAL_Evaluation_was_sie_bringen.pdf; 1.10.2019].

SEVAL – Schweizerische Evaluationsgesellschaft (2016): *Evaluationsstandards der Schweizerischen Evaluationsgesellschaft*. Fribourg: SEVAL. [https://www.seval.ch/app/uploads/2018/01/SEVAL-Standards-2016_d.pdf; 1.10.2019].

SEVAL – Schweizerische Evaluationsgesellschaft (2017a): *Erläuterungen zu den Evaluationsstandards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards)*. Fribourg: SEVAL. [https://www.seval.ch/app/uploads/2018/01/SEVAL-Standards_2016-Erläuterungen.pdf; 1.10.2019].

SEVAL – Schweizerische Evaluationsgesellschaft (2017b): *Strategie des SEVAL-Vorstands 2018–2021*. Fribourg: SEVAL. [<https://www.seval.ch/app/uploads/2018/08/Strategie2018-2021-def-d.pdf>; 1.10.2019].

SEVAL – Schweizerische Evaluationsgesellschaft (2018): *Qualitätsmanagement Lehrveranstaltungen: Empfehlungen für Konzipierung, Etablierung, Durchführung und Verwendung von Evaluationen*. [https://www.seval.ch/app/uploads/2017/10/Empfehlungen_Qualitaetsmanagement_Lehrveranstaltungen.pdf; 1.10.2019].

WIDMER, THOMAS (2011): *Zehn Jahre Evaluationsstandards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards)*. In: Schweizerische Zeitschrift für Kriminologie 10/2. S. 23–30.

WIDMER, THOMAS et al. (2016): *Forschung über Evaluation in der Schweiz: Stand und Aussichten*. In: LeGes – Gesetzgebung & Evaluation 27/3. S. 459–483.

WIDMER, THOMAS; BEYWL, WOLFGANG und FABIAN, CARLO (Hrsg.) (2009): *Evaluation – Ein systematisches Handbuch*. Wiesbaden: VS Verlag für Sozialwissenschaften.

WIDMER, THOMAS; LANDERT, CHARLES und BACHMANN, NICOLE (2001): *Evaluations-Standards der Schweizerischen Evaluationsgesellschaft*. [https://www.seval.ch/app/uploads/2017/06/SEVAL_Standards_2001_D.pdf; 1.10.2019].

WIDMER, THOMAS und DE ROCCHI, THOMAS (2012): *Evaluation: Grundlagen, Ansätze und Anwendungen*. Zürich: Rüegger.