

INTERNE EVALUIERUNG IM ÖSTERREICHISCHEN FORSCHUNGSBEREICH RÜCKSCHAU UND REFLEXION ZUR FTEVAL TAGUNG “VERBINDLICHKEIT IN DER INTERNEN EVALUIERUNG”

ISABELLA WAGNER, ELISABETH FROSCHAUER-NEUHAUSER AND KATHARINA WARTA
10.22163/fteval.2021.520

KURZFASSUNG

In diesem Beitrag wird die Diskussion der fteval Tagung “Verbindlichkeit in der internen Evaluierung”, die am 30. November 2020 online stattgefunden hat, eingefangen. Der analytische Rahmen wird vertieft und verschiedene Aspekte interner Evaluierung reflektiert. So wird interne Evaluierung als Prozess verstanden, unterschiedliche Schwierigkeiten nach ihrer Durchführung und ihrem Zentrum der Verantwortung (jeweils intern oder extern) betrachtet und diese Überlegungen in den Kontext der Beispielfälle der Tagung eingebettet.

ABSTRACT

With this contribution, we want to wrap up the discussions during the fteval symposium “Commitment in Internal Evaluations” that took place on-line on 30 November 2020. We present and further elaborate the analytical framework used and reflect on different aspects of internal evaluation. By understanding internal evaluation as a process, and observing differences depending on the main responsibility and way of execution (internal or external), we put our reflexions in context with the exemplary cases presented at the symposium.

EINLEITUNG

Im Bereich der Forschungs- Technologie- und Innovationspolitik (FTI-Politik) ist die Evaluierung von Programmen und Instrumenten gut etabliert. In Österreich legen allgemeine gesetzliche Erfordernisse, spezifische Anforderungen im Kontext von Richtlinien und Fördertätigkeiten sowie haushaltrechtliche Maßgaben die Durchführung von Evaluierungen fest. Wie die im Jahr 2019 im Auftrag der fteval durchgeführte Marktstudie zeigt, handelt es sich hierbei so gut wie ausschließlich um externe Programmevaluierungen, während Institutionen und Organisationen (aber auch gesetzliche Regelungen) eher selten evaluiert werden (Streicher et al., 2020). Aus Erfahrungsberichten ist bekannt, dass strukturierte interne Evaluierungen im Bereich der Forschungs- und Technologiepolitik selten sind, auch wenn insbesondere Förderagenturen durch-

aus interne Untersuchungen zur Wirksamkeit bestimmter Maßnahmen und Instrumente durchführen.

Was sind nun interne Evaluierungen? Die DeGEval definiert in ihren Standards eine Evaluation als intern, wenn die Evaluierenden derselben Organisation angehören, die auch für den Evaluationsgegenstand verantwortlich ist. Diese kann als Selbstevaluation oder „in-house Evaluation“¹ realisiert werden. Bei einer Selbstevaluation evaluieren die Personen, die einen Gegenstand entwickeln und/oder umsetzen, diesen selbstverantwortlich und eigenständig und sind sich häufig auch selbst Auftraggebende. Bei einer in-house Evaluation fallen Programmverantwortung und Evaluierende auseinander, wobei aber noch immer alle beteiligten Personen derselben Organisation angehören. Bei einer externen Evaluation sind die Evaluierenden nicht Teil derselben Organisation. Wir werden im Folgenden auf die genannten Differenzierungen und etwaige Überlappungen eingehen.

Interne Evaluation stellt eine strukturierte Reflexion und Auswertung zur Erfolgskontrolle der eigenen Aktivitäten ins Zentrum – in Bezug auf eigene oder die mit Auftraggebenden definierten Ziele, sowie zur Weiterentwicklung der eigenen Institution und ihrer Aktivitäten. Bei der Darstellung von Wirkungen und der Qualität der eigenen Arbeit hat eine interne Evaluierung im Vergleich zur externen Evaluierung Informationsvorteile, Aufgabe ist auch hier, dies strukturiert und nachvollziehbar darzustellen, nach innen oder außen (oder beides). Die Prozesse der Umsetzung der Empfehlungen sowie die Überprüfung dieser Umsetzung sind vielfältig. Sie hängen nicht zuletzt davon ab, ob Evaluierungsergebnisse oder die Umsetzung von Empfehlungen bzw. Auflagen in Verbindung mit Finanzierungszusagen stehen, oder ob eine Evaluierung in erster Linie formativ angelegt ist, also der Information und dem Lernen dient. Aspekte wie Objektivität, Validität, Reliabilität, Transparenz, Verlässlichkeit in Bezug auf die Umsetzung, sowie Verwertung und Reflexion der Ergebnisse stehen bei internen Evaluierungen jedenfalls auf dem Prüfstand.

In diesem Beitrag gehen wir der Frage nach, in welcher Weise es einen Unterschied macht, ob man sich bei einer internen Evaluation vor sich selbst oder einer Instanz von außen erklärt und wer die Evaluierung durchführt. Wir beziehen uns hierbei in erster Linie auf die Ergebnisse einer online-Veranstaltung, ergänzt durch Literatur sowie Erfahrungen aus dem universitären Hochschulsektor, in dem interne Evaluierungen einen anderen Stellenwert haben als in der Programm-basierten FTI-Förderung. Aus der anregenden Diskussion der fteval Tagung „Verbindlich-

keit in der internen Evaluierung“² schöpfend, entstand dieser Artikel als Rückschau, vertiefende Reflexion und in Bezugnahme auf den österreichischen Kontext. Das Autorinnen-Trio besteht aus der federführenden Veranstalterin an der fteval (Isabella Wagner), der ehemaligen Obfrau der fteval und Initiatorin der Tagung (Katharina Warta, Technopolis) sowie Elisabeth Froschauer-Neuhauser (AQ) als maßgeblich Mitwirkende an der Sonderausgabe des fteval Journals zu Hochschulevaluierung (Nr. 49). Im Rahmen der Tagung wurden gemeinsam mit ausgewählten Vortragenden der Alexander von Humboldt-Stiftung (AvH), des Instituts für Höhere Studien (IHS) und der Ludwig Boltzmann Gesellschaft (LBG) sowie Discussants der Universität Wien, der Christian Doppler Forschungsgesellschaft (CDG) und des Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF) die Besonderheiten interner Evaluation beleuchtet und Beispiele erfolgreicher Umsetzungen diskutiert. Folgende Themen leiteten den Erfahrungs- und Wissensaustausch:

Wert interner Evaluierung

- Spannungsfeld zwischen Objektivität und interner oder externer Vereinnahmung
- Spannungsfeld zwischen Lernen / Entwicklung und Rechtfertigung
- Zugänglichkeit von Daten und Ergebnissen, Ownership der Evaluierung, Transparenz
- Anonymisierung der Daten - Offenheit der Ergebnislegung
- Aufwand und Budget (personelle und finanzielle Ressourcen) von internen und externen Evaluationen

Risiken bzw. Herausforderungen der internen Evaluation und der Selbstevaluierung im Speziellen

- Unabhängigkeit und Unparteilichkeit der Evaluator*innen, oder blinde Flecken
- Abgrenzung zu sonstigen Reflexions- und Organisationsentwicklungsformaten
- Aufwand für die Evaluierung, insbesondere bei Selbstevaluierungen, Wahrnehmung von Evaluierung als Belastung statt als Unterstützung, Widerstände
- Unterschiede im Umgang mit Transparenz und etwaigen Machtverschiebungen und ihre Konsequenzen auf die Dynamik einer internen Evaluierung
- Verbindlichkeit in der Umsetzung der Empfehlungen
- Bevor wir auf konkrete Erfahrungen eingehen, stellen wir im kommenden Abschnitt ein Modell als Basis zur Differenzierung unterschiedlicher Evaluierungssettings vor.

AUSPRÄGUNGEN INTERNER EVALUIERUNG

Um ein gemeinsames Verständnis der verschiedenen Ausprägungen interner Evaluierung zu generieren, wurde eine zweidimensionale Matrix nach Durchführung und dem Schwerpunkt der Verantwortung erstellt, wobei beide Achsen jeweils die Tendenz beschreiben, ob die Aktivität/Verantwortung innerhalb oder außerhalb des Evaluationsobjekt liegt. Die horizontale Achse der Durchführung betrachtet, ob die Evaluation inner-

halb einer Institution durchgeführt wurde, oder (auch) durch externe Evaluator*innen bzw. durchführende Agenturen. Wie bereits erwähnt, wird bei der Durchführung innerhalb derselben Organisation nur dann von „Selbstevaluation“ gesprochen, wenn die Evaluierenden gleichzeitig für den Evaluierungsgegenstand zuständig sind.

Die vertikale Achse hat mehrere Bedeutungsebenen: sie stellt die Zielgruppe dar, die Frage, an wen sich die Evaluierungsergebnisse richten, wer das „Gegenüber“ der Evaluierung ist und wer letztendlich für die Umsetzung der Evaluierungsergebnisse verantwortlich ist. Diese Position kann mit der Beauftragung zusammenfallen, aber auch mit einer Instanz, gegenüber der es gilt, die eigenen Tätigkeiten darzustellen, oder gar rechtfertigen. Umgekehrt dient eine vorrangig nach innen gerichtete Evaluation eher dem organisationalen Lernen, die Berichte stehen dementsprechend kaum der Öffentlichkeit zur Verfügung. Diese Achse beschreibt also die Lage der Verantwortung im Spannungsfeld Evaluierungsobjekt, Durchführende und Beauftragende.

Die räumliche Darstellung verdeutlicht, dass es in der Regel Mischformen gibt. Es entstehen vier Quadranten (Abbildung 1), die jeweils unterschiedliche Ausprägungen beschreiben:

- **Durchführung und Verantwortung intern:** Bei interner Beauftragung und interner Durchführung spricht man von „Selbstevaluation“, wenn die Evaluierung durch die für den Evaluationsgegenstand zuständigen Personen durchgeführt wird. Ebenfalls intern, aber durch z.B.: eine andere Abteilung (z.B. Qualitätssicherung) durchgeführt, wird die Evaluierung als in-house Evaluierung bezeichnet.
- **Durchführung intern – Verantwortung extern:** Es gibt unterschiedliche Ausprägungen, wie externe Elemente in der internen Evaluierung eine Rolle spielen. So kann etwa die Beauftragung von einer Instanz außerhalb des Evaluationsgegenstandes kommen und die Durchführung intern abgewickelt werden. Das könnten etwa Monitoring-Aktivitäten sein, die einer auftraggebenden oder fördergebenden Institution vorzulegen sind, was beispielsweise an den österreichischen Universitäten im Zuge der individuellen Leistungsvereinbarungen mit dem Ministerium der Fall ist. Ebenso fallen Selbstevaluierungsberichte in diese Kategorie, die im Rahmen einer externen Evaluierung und nach Vorgaben, entlang von Fragen des externen Evaluierungsteams erstellt werden.
- **Durchführung extern – Verantwortung intern:** Die andere Variante, sich externer Akteur*innen zu bedienen, ist wenn zwar Eigeninitiative/Selbstbeauftragung die Evaluation initiiert, aber die Durchführung der Evaluierung (teilweise) an externe Expert*innen vergeben wird. Das ist beispielsweise der Fall, wenn die Daten intern gesammelt, aber ein externes Expert*innengremium zur Auswertung engagiert wird. Die Entscheidung über die Nutzung und Verbreitung der Evaluierungsergebnisse liegt in der Organisation, die für den Evaluierungsgegenstand zuständig ist („Ownership“).
- **Durchführung und Verantwortung extern:** Wenn sowohl die Verantwortung als auch die Durchführung außerhalb der evaluierten Institution liegen, spricht man von einer externen Evaluierung. Dennoch sind die Evaluationsobjekte aktiver Bestandteil der Evaluierung und die Konsequenzen betreffen immer auch

interne Prozesse. Die Organisation hat in der Regel Daten bereitzustellen und ist dadurch mit einem gewissen Aufwand konfrontiert. Viele Programmevaluationen fallen in diese Kategorie. An den österreichischen Universitäten und Fachhochschulen zählen z.B. auch die Audits dazu, wo der Fördergeber (das Ministerium) per Gesetz vorgegeben hat, dass das interne Qualitätsmanagement regelmäßig einer Evaluation durch externe Agenturen zu unterziehen ist.

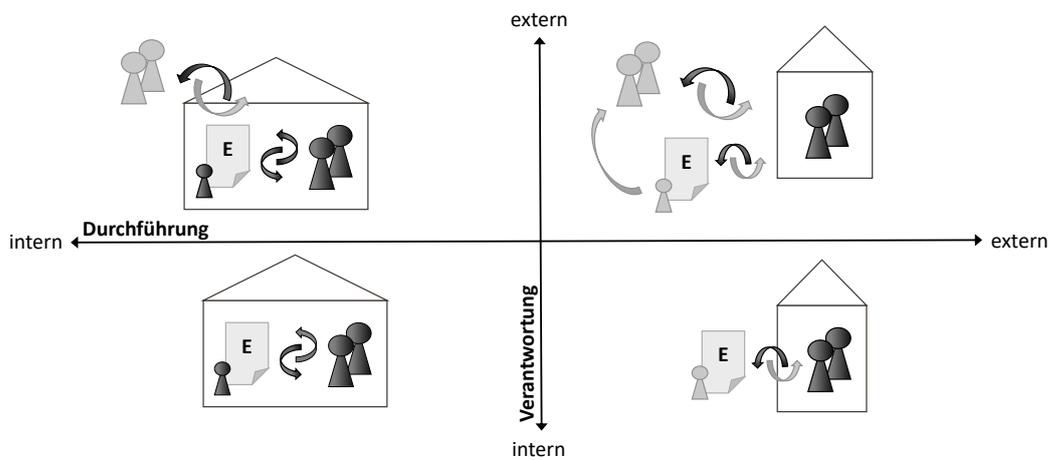
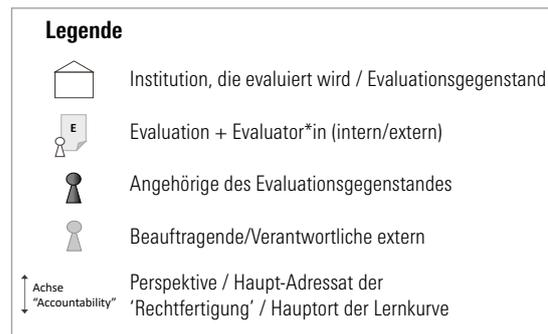


Abbildung 1: Matrix interner Evaluierungs-Zugänge nach Durchführung (intern/extern) und Verantwortung (intern/extern) - am Beispiel einer Institution als Evaluationsobjekt

In unserer Betrachtung wird nun anhand von unterschiedlichen Beispielen beschrieben, wie sich die Verortung einer Evaluierung entlang der beiden Achsen jeweils auf Ergebnisse und Empfehlungen bzw. deren Umsetzung auswirken kann.

Weitere Beispiele aus der Praxis des universitären Hochschulsektors werden eingebracht um unsere Überlegungen eingänglicher zu machen und breiter zu verorten. Konkret wurden die Technische Universität Graz (TU Graz) und die Universität für angewandte Kunst Wien (die Angewandte) als Fälle beschrieben.

AUSWAHL DER BEISPIELE

Die Auswahl der Beispiele für die fteval Tagung hatte den Anspruch einer großen Diversität der Zugänge zu interner Evaluierung im Sinne der beschriebenen Achsen. So stellte Meike Olbrecht, Leiterin des Referats für Evaluation und Statistik an der Alexander von Humboldt-Stiftung, deren Modell vor, in der die Beauftragung zwar extern verortet ist, Verantwortung und Durchführung aber intern in Form einer in-house Evaluierung durch die Stabsstelle Evaluation & Statistik stattfindet. Im Beispiel des Instituts für Höhere Studien stellte Thomas König eine rezente interne Evaluierung mit Hilfe eines externen Expert*innen-Gremiums vor, die zwar aus sich selbst heraus beauftragt war, aber gleichzeitig zur externen Darstellung, wie auch als Grundlage zur institutionellen Weiterentwicklung diente. Im Fall der Ludwig Boltzmann Gesellschaft beschrieb Patrick Lehner (Open Innovation in Science Center - OIS), wie sowohl die Beauftragung als auch die Durchführung in seiner Organisation derzeit internen Logiken folgt, wobei sich das im Zuge anstehender Umgestaltungen ändern könnte.

EVALUIERUNG, BERATUNG, REFLEXION – EINE ABGRENZUNG

Der Begriff „Evaluierung“ wird in unterschiedlichen Kontexten sehr breit und beinahe großzügig verwendet. So steht die strategische Orientierung als „systematische Untersuchung der Güte oder des Nutzens eines Evaluierungsgegenstandes“ (ZIT DeGEval Standards, siehe fteval Standards, Kohlweg, 2019) im Sinne der Ermöglichung von Lern- und Verbesserungsschleifen der Wahrnehmung verkürzter Kontrolle durch Dritte gegenüber. Eine klare Abgrenzung zu anderen Formaten von Reflexionsprozessen wird kaum unternommen und stellt sich bei Betrachtung der einzelnen Verfahrenstypen auch oft als schwierig heraus. Gerade bei internen Evaluierungen lohnt es sich, sich vorab stärker mit der Zielsetzung auseinanderzusetzen.

Prozessorientierte Methoden, die gruppaes und organisationales Lernen durch Feedback und Reflexion unterstützen, gehen davon aus, dass es schwierig ist, sich mit eigenen Schwächen auseinanderzusetzen und dass es hierfür einen möglichst sicheren Rahmen braucht, der es den Teilnehmenden erlaubt, sich zu öffnen und neue Fragen und Perspektiven zuzulassen. Die Bewertung wird hintenangestellt, es geht weniger um quantitative Erfolgsindikatoren als um Entwicklungspotenziale. Evaluierung als Verbindung von „evidence and judgement“ vertritt demgegenüber einen objektivierten Standpunkt. Bei internen Evaluierungen kann nun der „Raum“, in dem beobachtet und bewertet wird, mit dem Raum zusammenfallen, der mit der Bewertung zurecht kommen soll. Dies kann sich schwierig gestalten, wenn Fragen von Macht, Zugehörigkeit aber auch persönlicher Nähe ins Spiel kommen. So macht es einen großen Unterschied, ob im Rahmen einer internen Evaluierung zu erwarten ist, dass ein negatives Ergebnis direkte Konsequenzen auf zukünftige Arbeitsbereiche oder finanzielle Ressourcen hat – sollte diese Sorge bestehen, dann kehren sich die Vorteile der Vertrautheit und Offenheit möglicherweise ins Gegenteil um, und es wird im Vergleich zu externen Evaluierungen schwieriger, transparent und objektiv zu argumentieren. Je stärker materielle Konsequenzen oder eine Einflussnahme durch Dritte zu erwarten sind, desto stärker sind die Anreize, die Selbstevaluierung in die für sich gewünschte Richtung zu bewegen. Wenn hingegen keine Gefahr der Einflussnahmen von Außen besteht, dann ist interne Evaluierung – als eine evidenz-basierte Auswertung – ein geeignetes Instrument interner Reflexion.

An dieser Stelle sei noch angemerkt, dass die drei Prozesse externe Evaluierung, Selbstevaluierung und Reflexion häufig in Kombination stattfinden. So kann beispielsweise eine extern angelegte Evaluierung einer Institution einen Selbstevaluierungsbericht dieser Institution als eine der Informationsquellen beinhalten. Dieser Bericht ist häufig um eine positive Darstellung bemüht – löst aber gleichzeitig im Zuge der Erstellung Reflexionsprozesse in der Organisation aus.

Die fteval Standards nennen 12 Grundprinzipien für FTI-Evaluierungen (fteval, 2019). Wir halten die Betrachtung dieser Standards für hilfreich, um in Abgrenzung zu anderen Prozessen wie Reflexion, Monitoring, Steuerung und Organisationsentwicklung zu überlegen, was einen internen Bewertungsprozess zu einer „Evaluierung“ macht. Zwei der Prinzipien sollen hervorgehoben werden:

- **„Unparteilichkeit“** setzt voraus, dass „der gesamte Evaluierungsprozess und die Ergebnisse auf unvoreingenommene Handlungen und Positionen schließen lassen und nachvollziehbar und transparent zustande gekommen sind“ (S. 16). Sobald in Organisationen Hierarchien, Verbindlichkeiten, Näheverhältnisse und Abhängigkeiten existieren, ist auf diese Unparteilichkeit – die auch bei externen Evaluierungen zu sichern ist – besonders zu achten.

Um diese Unparteilichkeit zu gewährleisten, hat es sich an den Universitäten in den letzten Jahren etabliert, Stabstellen oder Referate für Evaluierung einzurichten, die direkt unter den Rektoraten angeordnet sind, somit keinen Fachrichtungen zugeordnet sind und möglichst eigenständig agieren. Durch Abgrenzung zum Rektorat wird gleichzeitig auch die Akzeptanz der Maßnahmen erhöht, da sie dadurch weniger als „vom direkten Vorgesetzten verordnet“ wahrgenommen werden.

Bei der Alexander von Humboldt-Stiftung (AvH) sind vier Akteure in interne Evaluierungen involviert: Die Stabsstelle für Evaluation und Statistik (E&S) wurde vor 15 Jahren eingerichtet und ist – neben der Betreuung externer Evaluierungen – in Zusammenarbeit mit stiftungsinternen Projektteams für die Erstellung interner Evaluierungen zuständig. Der AvH steht weiter ein externer wissenschaftlicher Beirat zur Seite; schließlich entscheiden die geldgebenden Bundesministerien und der Stiftungsrat über den Evaluierungszeitplan. E&S Mitarbeiter*innen haben keine Programmverantwortung und sehen ihre Aufgabe auch bei internen Evaluierungen in kollegialer Beratung und der Förderung des Austauschs. Die Unparteilichkeit der Stabsstelle setzt voraus, dass organisational ausreichend Distanz besteht, dass die Evaluationsaufgaben und -ziele eindeutig definiert sind und dass die Rollen aller Beteiligten zu Beginn geklärt werden. Transparenz und Vertrauen sind eine Grundlage für die Nutzung als auch den Nutzen der internen Evaluierung. Die Diskussion der Ergebnisse, Schlussfolgerungen und Empfehlungen mit „Dritten“, hier dem Beirat für Programmevaluationen, erhöht die Objektivität und stärkt letztendlich die Rolle der Stabsstelle. Im Beispiel des IHS wurde ein externes Sekretariat zur Abwicklung der sowohl intern als auch mit externen Evaluierenden durchgeführten Studie eingerichtet. Die Interpretation der Ergebnisse erfolgte allerdings intern. Am OIS Center der LBG versucht man in ihrer aktuellen Ausgestaltung die Wirksamkeit der eigenen Aktivitäten hinsichtlich der Ziele des Centers laufend zu überprüfen. Hierzu wurde ein Zugang basierend auf einer umfassenden „Theory of Change“ gewählt, der auch Experimentierraum zulässt. Inwiefern dieser Zugang allerdings auch Überparteilichkeit garantieren kann, bleibt dabei unklar, trotz deutlicher Bemühungen zur Transparenz.³

- Das Prinzip **„Unabhängigkeit“** wird in den Standards wie folgt definiert: „Eine Evaluierung wird mithilfe von wissenschaftlichen Methoden erstellt und unterschiedliche Meinungen bzw. Positionen sind berücksichtigt. Die Evaluierung ist weder von Politik, Auftraggeber*innen, dem Programmmanagement, den Betroffenen oder auch von einem möglichen Bias der Evaluator*innen selbst, inhaltlich beeinflusst bzw. manipuliert. (...)“ (S. 16).

Aus diesem Grund wird externen Evaluationsprozessen häufig eine höhere Wertigkeit zugeschrieben. Zum einen wird durch die Beauftragung einer externen Agentur die Unparteilichkeit sowohl in Richtung Auftraggeber*in als auch der zu evaluierenden Einheit gewährleistet. Dies wird nochmals durch die Auswahl der Peers unterstützt. So wird z.B. bei Evaluierungen von Fakultäten an der TU Graz oder der Angewandten die Durchführung durch eine externe Agentur ganz stark mit den Begriffen der Unabhängigkeit und der Transparenz verknüpft (Gaberscik & Berner, 2019 und Kernegger, 2019). Im Fall von Selbstevaluierungen ist die Lage anders: Hier kann man mittels der Auswahl einer geeigneten Methode ausreichend Objektivität sicherstellen. Die Herausforderung liegt vermutlich darin, die geeigneten Methoden und Akteure zu wählen, die in die Evaluierung eingebunden werden. Es kann bei einer Selbstevaluierung jedoch

nicht von einem "Gutachten" die Rede sein, hier befindet man sich ganz klar im Bereich der Selbstreflexion. An den Beispielen der Evaluation des IHS und der Praxis der AvH sieht man den Versuch, durch externe Elemente sowie der Art ihrer Einbettung diese Unabhängigkeit zu erzeugen. So wirken an der AvH sowohl eine Stabstelle für Evaluierung, ein externer Beirat, als auch interne Projektteams und der Stiftungsrat als Akteur*innen gemeinsamer interner Evaluationsvorhaben. Das IHS zog wie erwähnt interne wie externe Expertise zu Rate, überließ die Abwicklung dabei aber klar einem externen Sekretariat. Die LBG möchte bei der regelmäßigen und systematischen Evaluierung ihrer Ludwig-Boltzmann Institute (LBIs) die Unabhängigkeit der Vorgangsweise nicht nur durch das Zuziehen externer, sondern sogar ausschließlich internationaler Expert*innen sicherstellen.

BEAUFTRAGUNG UND EVALUATIONSOBJEKT

Die Beauftragung spielt bei Evaluierungen eine wesentliche Rolle: Im Bereich der FTI-Politik geht es um die Verwendung öffentlicher Gelder. Der Bedarf an Evaluierung kann sich entweder aus rechtlichen Bestimmungen ergeben, oder aus einem reinen Informationsbedürfnis. Zur Beauftragung einer Evaluierung gehören mehrere Schritte: Die Definition des Evaluationsgegenstandes, die Definition von Evaluierungsfragen, gegebenenfalls die Auswahl von Expert*innen zur Durchführung, ein gewisses Maß an Qualitätskontrolle, der offizielle Abschluss des Evaluierungsprojekts durch die Abnahme eines Endberichts, gegebenenfalls die Freigabe zur Veröffentlichung, jedenfalls die Kommunikation der Ergebnisse an relevante Zielgruppen, in der Folge gegebenenfalls auch die Verantwortung für die Umsetzung von Empfehlungen, bzw. die Beaufsichtigung derselben. Für eine Evaluierung ist es erfahrungsgemäß gut, wenn es ein starkes "Ownership" bei gleichzeitig gesicherter Unabhängigkeit der Evaluierung gibt. Dieses Ownership liegt entweder in der externen Governance (Politik, Programmeigentümer, Aufsichtsorgan), in der internen Leitung (Geschäftsführung, sonstige Leitungsorgane), in einer Stabstelle oder Abteilung für Evaluierung, Monitoring oder Qualitätssicherung oder aber bei den Verantwortlichen für den Evaluierungsgegenstand selbst. Wir halten es in jedem dieser Fälle für unabdingbar, dass zu Beginn des Evaluierungsvorhabens geklärt wird, wer "Eigentümer*in" der Ergebnisse ist, und wer über diese Ergebnisse jedenfalls informiert wird, da dies die Verbindlichkeit des Evaluierungsvorhabens erhöht.

Auch hier kann sich die Einbindung externer Agenturen als vorteilhaft herausstellen. Die Erfahrung hat gezeigt, dass Auftraggeber*innen zwar das Ziel einer Evaluierung nennen können, sich aber über Verantwortlichkeiten und die Wichtigkeit von klaren Definitionen meist nicht bewusst sind (Froschauer-Neuhauser et al., 2020).

Dies ist bei begleitenden Evaluierungen eine besondere Herausforderung, da hier häufig die Programmverantwortlichen Auftraggeber*innen sind, aber angesichts der Unsicherheit, in welche Richtung die Zwischenergebnisse der Evaluierung ausschlagen, nicht festgelegt wird, mit wem die Ergebnisse geteilt werden. Möglicherweise wollen sich Auftraggeber*innen vorbehalten, ohne zu viel Öffentlichkeit an diesen Ergebnissen zu lernen. Wenn jedoch im Prozess eine Triangulation fehlt, hat dies einen Einfluss auf die Fortsetzung des Evaluierungsprozesses,

der dann möglicherweise in Richtung Prozessberatung und -begleitung tendiert. Dies mag dem Bedarf entsprechen und durchaus nützlich sein, es haben sich jedoch im Feld von Coaching und Prozessbegleitung teilweise andere Qualitätsstandards etabliert als bei der Evaluierung.

Meike Olbrecht von der AvH illustrierte am Beispiel der Evaluierung eines großen Stipendienprogramms der Stiftung die unterschiedlichen Ausrichtungen der externen und der internen Evaluierung: Jedes Programm der AvH wird alle 10 Jahre durch externe Dienstleistende hinsichtlich der Erreichung der Programmziele evaluiert. Die Berichte werden im wissenschaftlichen Beirat und dem Stiftungsrat diskutiert und dann auf der Internetseite der Stiftung – gegebenenfalls mit einer Stellungnahme zu den Empfehlungen – veröffentlicht. Fünf Jahre nach der externen Evaluierung führt das Referat „E&S“ eine interne Evaluierung der Umsetzung der Empfehlungen durch und erstellt auf dieser Basis einen Zwischenbericht. Zusätzlich informiert dieser Bericht über statistische Trends und aktuelle Kennziffern der Programmentwicklung. Auch dieser wird im wissenschaftlichen Beirat und in der Regel im Stiftungsrat diskutiert, und dann an die Geldgebenden gesendet, nicht jedoch veröffentlicht. Die Ergebnisse dienen zum einen internen Anpassungen, zum anderen der Vorbereitung der nächsten Programmevaluierung.

DIE KONKRETE UMSETZUNG (INTERN/EXTERN)

Neben der Definition der Eckpunkte der Verantwortung über die Ergebnisse braucht es für die konkrete Umsetzung einer Evaluation eine genaue Definition darüber, was es herauszufinden gilt. Der Evaluationsinhalt und die benötigten Kompetenzen für die Evaluierung müssen festgestellt werden. Aufgrund dieser Definitionen kann man sich im Anschluss für eine passende Form der Umsetzung (extern/intern/Mischform) entscheiden.

Im Hochschulsektor haben Evaluierungen in den letzten Jahren stark zugenommen, nicht zuletzt dadurch, dass sie von Seiten des Ministeriums für Bildung, Wissenschaft und Forschung zu verschiedenen Themen im Rahmen der Leistungsvereinbarungen eingefordert werden. Viele Universitäten versuchen die Rechenschaftslegung aber auch möglichst effektiv mit der Möglichkeit zur internen Weiterentwicklung zu verknüpfen. Der Trend geht dahin, Evaluierungen intern anzustoßen und so zu gestalten, dass man sie auch für die Darstellung und die Rechtfertigung nach Außen verwenden kann. Damit einhergehend ist eine stärkere Systematisierung und Formalisierung der Evaluierungen – die Stabstellen für Qualitätssicherung und Evaluierungen haben Konzepte zu regelmäßigen Evaluierungen von Einheiten entwickelt, die zudem auch transparent an alle Universitätsangehörigen kommuniziert sind.

Die TU Graz sieht beispielsweise eine Evaluierung ihrer Fakultäten in einem 7-Jahreszyklus vor. Das bedeutet, dass jedes Jahr eine der Fakultäten evaluiert wird. Hierbei ist der Prozess von der Ausschreibung bis hin zu den vordefinierten Kernthemen formalisiert. Gleichzeitig erhält man sich die Flexibilität auf aktuelle Themen von außen reagieren zu können – nachdem das Ministerium in den Leistungsvereinbarungen mit den Universitäten eine starke Zielsetzung auf „Studierbarkeit“ definiert hat, wird dieses Thema auch stärker in den laufenden Evaluationen mitaufgenommen. Individualität findet in der jeweiligen Ausgestaltung Platz – so sind unterschiedliche Fokus-Setzungen möglich oder es wird auch ein Mitspracherecht bei der Peer-Auswahl eingeräumt.

Kommen wir noch einmal auf die AvH zurück: Die Schnittstelle zu externen Evaluator*innen ist zum einen eine öffentliche Ausschreibung mit Pflichtenheft in einem zweistufigen Verfahren (schriftliches Angebot, Vortrag und Diskussion des Angebots mit dem Evaluierungsbeirat), gefolgt von der unabhängigen Evaluierung entlang der formulierten Evaluierungsfragen durch externe Expert*innen und die Abnahme des Endberichts durch die Geschäftsführung der AvH auf Empfehlung des Beirats. Die interne Evaluierung orientiert sich an den Empfehlungen, dokumentiert die Initiativen der Umsetzung und betrachtet dann mittels ergänzender Untersuchungen den Erfolg dieser Umsetzung. So wurde beispielsweise die Empfehlung aus einer externen Programmevaluation zum Humboldt Forschungsstipendien-Programm aus dem Jahr 2011, unterrepräsentierte Länder im Stipendienprogramm stärker einzubinden, in zwei Schritten umgesetzt: Aufbauend auf einer internen Potenzialanalyse wählte die AvH fünf Schwerpunktländer und führte in den Jahren 2012-2018 eine Schwerpunktländerinitiative durch, indem sie Maßnahmen zur Erhöhung der Anzahl von Bewerbungen und Bewilligungen in diesen Ländern durchführte, wodurch diese in allen Ländern gesteigert werden konnten. Solche Ergebnisse, aber auch weniger erfolgreiche Umsetzungsversuche, werden im Zwischenbericht dokumentiert.

WIE WIRD NUN VERBINDLICHKEIT ERZEUGT? DIE ANWENDUNG DER ERGEBNISSE UND DER NUTZEN INTERNER EVALUIERUNG

Jede Evaluierung, egal ob extern, intern oder in einer Mischform, hat einen Effekt auf das Evaluationsobjekt. Auch in Fällen, in denen die Accountability gegenüber unterschiedlichen Stakeholdern im Zentrum steht, ergibt sich durch die Reflexion des eigenen Tuns bzw. durch die vertiefte Auseinandersetzung mit den eigenen Key Performance Indicators (KPIs) auch immer ein Effekt nach innen. Je stärker es Richtung „intern“ geht, desto stärker legt sich die Gewichtung auf die Weiterentwicklung des Evaluationsobjekts.

Die Umsetzung von Schlussfolgerungen und Empfehlungen einer Evaluierung kann sowohl davon abhängen, wie viel Druck erzeugt werden kann, als auch, wieviel Verständnis für das Verbesserungspotenzial erzeugt wird. So gibt es einige Förderprogramme, bei denen Evaluierungspanels die Möglichkeit haben, die Förderung an die Umsetzung bestimmter Auflagen zu knüpfen. Dies erzeugt eine hohe Verbindlichkeit. Häufig ist es aber auch ein Anliegen, dass die dargestellten Ergebnisse ausreichend überzeugend und nachvollziehbar sind, damit Entscheidungsträger*innen die Empfehlungen aufgreifen. Interessanterweise hat in einer dynamischen, an der (Organisations-)Entwicklung ausgerichteten Sichtweise die sogenannte „Gutachter*innenberatung“ keine gute Nachrede: „Bis in die Siebzigerjahre des vorigen Jahrhunderts war sie die vorherrschende Form der Unternehmensberatung, bei der externe oder interne Expertinnen und Experten bzw. Stabsleute eine Organisation untersuchten und den Auftraggebenden ihren Analysebefund mit Vorschlägen für Verbesserungen vorlegten. Die Beratung war mit dem „Gutachten“ für empfohlene Veränderungen bereits erledigt, denn die

Umsetzung wurde als selbstverständliche Sache des Managements gesehen. Und genau dort begannen für die Organisationen die eigentlichen Probleme.“ (Glasl et al, 2005: S. 34).

Es lässt sich beobachten, dass sich auch in der Evaluierungs-Community in den vergangenen Jahren etwas in der Herangehensweise verändert hat: Durch die Anerkennung unterschiedlicher partizipativer und gruppaler Erhebungs- und Analysemethoden neben der Analyse von Daten und Fakten sind auch Prozesse in den Fokus von Evaluierungen gerückt. So erweisen sich Besprechungen von Zwischenergebnissen, bei denen unterschiedliche Stakeholder eingeladen sind, sich mit den Beobachtungen und Auswertungen auseinanderzusetzen, als förderlich für eine Umsetzung späterer Empfehlungen. Es wächst also die Bereitschaft, prozessorientierte Vorgangsweisen (wie machen wir das?) mit möglichst objektiven Erfolgsindikatoren (wie viel ist gelungen?) zu kombinieren. Letztendlich gilt auch für interne Evaluierungen, dass sie sich an Nachvollziehbarkeit und Transparenz ausrichtet, um relevante Schlussfolgerungen und deren Umsetzung zu unterstützen. Das Beispiel der AvH zeigt, dass durch die Kombination eines sehr regelmäßigen Evaluierungsplans für alle Programme, ein gut ausgestattetes internes Referat für Evaluation und Statistik, sowie die Rückkopplung mit einem wissenschaftlichen Evaluationsbeirat ein hohes Maß an Verbindlichkeit erzeugt werden kann.

An der TU Graz wird die Verbindlichkeit der Evaluierungen insofern hergestellt, als dass die Fakultät nach Abschluss der Evaluierung gefordert ist, anhand der Empfehlungen zur Weiterentwicklung aus dem Peer-Bericht Arbeitsschritte abzuleiten, die auch mit konkreten Zeitplänen hinterlegt sind. Dieser Schritt erfolgt gesondert von der extern beauftragten Evaluierung nur noch universitätsintern. Sie stellen in der Folge auch die Grundlage für die jährlichen Folgegespräche zwischen Rektorat und Fakultätsleitung dar.

Die LBG verknüpft die Evaluierung der Ludwig Boltzmann-Institute mit der sehr konkreten Entscheidung, ob ein Institut weitergeführt werden soll oder nicht, was die Verbindlichkeit klar definiert. Im Fall der internen Reflexion im OIS Center ist die Verbindlichkeit nicht deutlich zu erfahren. Im Beispiel der IHS-Evaluierung war die Verbindlichkeit im deutlichen Wunsch zur Neuaufstellung und der Argumentation gegenüber dem Auftraggebenden verankert, somit wurden den Empfehlungen entsprechend die Zahl der Forschungsgruppen reduziert.

STIMMEN AUS DER FTEVAL COMMUNITY

Ausgangspunkt dieses Artikels war die fteval Tagung zu interner Evaluierung, bei der wir die Diskussionen der Teilnehmenden eingefangen haben. Diese möchten wir hier kurz darstellen.

Die rund 60 Akteur*innen, waren mehrheitlich selbst bereits Teil eines (internen) Evaluierungs-Vorhabens. Bei der Bewertung, wie stark oder wenig die Ergebnisse ernst genommen und umgesetzt wurden, fielen die meisten Antworten der Teilnehmenden im Mittelfeld aus.

In der mit „Mentimeter“ durchgeführten Umfrage waren die Teilnehmenden mehrheitlich der Überzeugung, dass in ihrer Institution derzeit eine interne Evaluierung mit Hilfe von externen Expert*innen am meisten Nutzen bringen würde. Eine rein interne oder rein extern durchgeführte Evaluierung wurde aktuell für die Heimatinstitution weniger nützlich eingeschätzt. Das deckt sich mit der allgemeinen Beobachtung,

dass diese Mischform auch die am häufigsten in der Praxis angewendete Variante ist.

Die Teilnehmenden wurden auch befragt, welche Dynamiken während interner Evaluierungen zu beobachten sind. Die Antworten sind in der folgenden Tabelle thematisch gruppiert dargestellt:

Welche Dynamiken sind bei internen Evaluierungen zu beobachten?	
Aspekt	Genannte Antworten
Vorbereitung und Ausrichtung	<ul style="list-style-type: none"> • Vereinbarung von Fokusthemen • Wirkung und Nutzen/Nutzung, Verhältnis von Aufwand und Ertrag • schwierige Einigung auf Prioritäten der Evaluation • Interne Evaluation von allen (intern) gewollt? • keine hohe Priorisierung (Zeit und Ressourcen)
Bereitschaft	<ul style="list-style-type: none"> • Hohe Bereitschaft Aspekte offen zu reflektieren • Anstoß für Veränderungen, die sonst kaum/schwer in Angriff genommen worden wären; Neuer Blick auf interne Prozesse und Abläufe
Widerstände	<ul style="list-style-type: none"> • Widerstände; Misstrauen; Skepsis; Umgang & Überwindung von Widerständen • Unwille; Fehlender Wille, wirklich aus der Evaluation zu lernen • Ängste; Angst vor Schelte von höherer Ebene, weil gewisse Kennzahlen nicht erfüllt werden, die die diverse Realität nur unzureichend wiedergeben • Verweigerung der Management Response • wenig Kooperationsbereitschaft; Kooperationsschwierigkeiten
Konflikt-potential	<ul style="list-style-type: none"> • Rollenkonflikt • Schwierige Position des Akteurs; „Minenfeld“; Diskreditierung der Evaluators*innen; Anzweifeln der Methodenkompetenz durch Management; Abwertung der internen Verfahren; mangelndes Vertrauen in die Fähigkeiten der internen Evaluators*innen • Interne Konkurrenzkämpfe; Positionskämpfe im Top- und Bereichsmanagement • Unterschiedliche Sichtweisen darauf, was wichtig ist Verhandlungen verschiedener Interessenslagen; unterschiedliche Prioritäten • Zutagetreten eines unterschiedlichen Verständnisses von internen Prozessen und Verantwortlichkeiten
Probleme im Prozess	<ul style="list-style-type: none"> • Rechtfertigung der Personen, die das Evaluierungsobjekt umgesetzt haben • Vertrauen in die Institution; Anonymisierung schwierig Infragestellen der Unabhängigkeit; fehlende Transparenz • Mangelnde Unabhängigkeit • Stress • fehlende/mangelnde Quality Culture • Verschiebung der Prioritäten und Zielsetzungen der Evaluation während diese bereits läuft
Ergebnisse	<ul style="list-style-type: none"> • Misstrauen, wie Ergebnisse verwendet werden; Follow up ist wichtig • fehlende/mangelnde „Quality Culture“
Ideen, Vor-/Ratschläge	<ul style="list-style-type: none"> • Stärkerer Fokus auf Vorbereitung und Umsetzungsmonitoring • Weg von der Rechenschaftslegung hin zur Profilentwicklung
Kritik	<ul style="list-style-type: none"> • Interne Evaluierungsergebnisse als Nabelschau • „Evaluitis“

Nach einer Diskussion darüber, welche Mechanismen zu installieren sind, damit eine interne Evaluation „garantiert scheitert“, wurden die Einschätzungen der Teilnehmenden dazu gesammelt – nun als „No-Gos“ formuliert. Die Antworten erstrecken sich über verschiedene Aspekte der Konzeption, Durchführung und Ergebnisbehandlung interner Evaluation. Sie sind in der folgenden Tabelle gelistet:

„No-Gos“ interner Evaluierung	
Aspekt	Genannte Antworten
Konzeption	<ul style="list-style-type: none"> • Kein echter Auftrag & keine interne Unterstützung; Unklarer Evaluationsauftrag; Unklarheit über Ziele und Konsequenzen • Zu ambitionierte / große Evaluierungsgegenstände & Ziele müssen auch erreichbar sein; unrealistische Evaluationsziele • ungenügende Ressourcen • Fehlendes Evaluationskonzept; Kein Evaluationskonzept • nur Vorgaben an die zu Evaluierenden, keine Gestaltungsmöglichkeiten, kein Raum zur Selbstreflexion • Hidden Agenda; Intransparenz; Fehlende Transparenz; die Evaluation „geheim“ zu halten • Nicht alle bei den Terms of References einbinden • Überlagerung mit anderen Evaluationen
Durchführung	<ul style="list-style-type: none"> • Inkompetente Personen beauftragen ohne methodisches Wissen • Unklare Rollen; Überlagerung von Prozessen (Evaluationen, Audits, ...) • Interne Konkurrenzkämpfe; Zweifel über Objektivität schüren; Konflikte nicht auflösen • die Beteiligten nicht einzubeziehen; Führungsgremium nicht einbinden; gewisse Stakeholder-Gruppen nicht evaluieren (z. B. befragen) • Mangelnder Umsetzungswillen • Kommunikation zwischen Beteiligten ist sehr wichtig; Schlechte Kommunikation
Ergebnisse	<ul style="list-style-type: none"> • Ergebnisse nicht/schlecht kommunizieren • sich nicht überlegen, was man mit Ergebnissen machen könnte; offenlassen, wie mit Empfehlungen umgegangen wird

CONCLUSIO

Die fteval Tagung „Verbindlichkeit in der internen Evaluierung“ hat ein großes Reflexionspotential und eine breite Vielfalt der Zugänge zu Evaluierung in der österreichischen Landschaft aufgezeigt. Mit Hilfe dieses Artikels haben wir anhand von Beispielen versucht, die unterschiedlichen Ausprägungen und deren Auswirkungen auf die Unabhängigkeit, das Ownership und die Verbindlichkeit in der Umsetzung zu diskutieren und sind zu folgenden zentralen Erkenntnissen gelangt:

- Interne Evaluierungen nehmen insgesamt zu, während sich die Art der Nutzung mitunter verschiebt. So sieht man beispielsweise an den Hochschulen mittlerweile stärker den Nutzen von internen Evaluationen in Hinblick auf die interne Weiterentwicklung und nicht allein als Darstellung gegenüber den Fördergebenden. Nicht ausschließlich die Rechenschaftslegung einem Beauftragenden gegenüber ist Ziel, sondern es werden interne Evaluationsansätze zur gesellschaftlichen Wirkungsmessung der eigenen Arbeit oder auch zur institutionellen Selbstreflexion der eigenen Organisationsentwicklung und somit auch der Accountability der eigenen Institutionsangehörigen gegenüber konzipiert.
- Die verschiedenen Ausprägungen von Evaluierung lassen sich nur bedingt abgrenzen, die Übergänge sind fließend. So wurden in den beobachteten Beispielen viele Mischformen von interner oder externer Durchführung gleichwohl wie interne und externe Verantwortung festgestellt. Es ist allerdings wichtig mitzudenken, welche Auswirkungen bestimmte Arten der Evaluation

auch in anderen Dimensionen haben können, wo beispielsweise Verantwortung hinsichtlich der Ergebnisverortung liegt

- Die Definition der Verantwortung ist für den Erfolg der Evaluierung und somit die Umsetzung der Empfehlungen sehr wichtig. Verbindlichkeit wurde in den Beispielen insbesondere durch eine klare räumliche und zeitliche Verankerung der Evaluierung in die Abläufe und Systeme einer Institution erreicht (festgelegte Evaluationsperioden bzw. eine eigene Einheit, ein Sekretariat oder eine Stabstelle für die Evaluationsvorhaben).
- Unabhängigkeit wird meist durch die Beziehung zu externen Akteur*innen sichergestellt. So die Funktionen transparent definiert und abgegrenzt sind, kann auch eine organisationsinterne Stabstelle einen gewissen Grad an Unabhängigkeit repräsentieren. Klare und transparente Kommunikation zwischen allen Beteiligten ist jedenfalls förderlich.

Wir hoffen mit diesem Beitrag die regen Diskussionen der Tagung eingefangen und verankert zu haben, um Impulse für weitere interne Evaluationsvorhaben zu geben. Bei der Tagung wurde der hohe personelle Aufwand als Schwierigkeit interner Evaluierungen hervorgehoben. Eine hohe Verbindlichkeit in der anschließenden Umsetzung der Ergebnisse kann diesen Aufwand lohnend machen.

REFERENZEN

Froschauer-Neuhauser, E., Kastelliz, D., Meznik, M. (2020). Gute Evaluation braucht klare Kommunikation – von Anfang an. Praxisbefunde aus dem Bereich hochschulischer Evaluation. 23. Jahrestagung der DeGEval. https://www.degeval.org/fileadmin/jahrestagung/Online_2020/Dokumentation/B4_Praesentation_AQAustria_DeGEval_2020.pdf

Gaberscik, G. und Berner, M. (2019). Evaluationsverfahren als Teil des universitären Qualitätsmanagements – Erfahrungen aus der Praxis der TU Graz. *fteval Journal for Research and Technology Policy Evaluation* (49), pp. 33-37. ISSN 1726-6629. <http://repository.fteval.at/501/>

Glasl, F., Kalcher, T., Piber, H. (Hrsg.). Professionelle Prozessberatung. Haupt Verlag, Bern, 2005

Österreichische Plattform für Forschungs- und Technologiepolitik-evaluierung (2019). Standards der Evaluierung in der Forschungs-, Technologie-, und Innovationspolitik. Technischer Bericht. *fteval - Österreichische Plattform für Forschungs- und Technologiepolitikevaluierung*. Wien. <https://repository.fteval.at/386/>

Kernegger, B. (2019). Evaluationsverfahren als Teil des universitären Qualitätsmanagements – Erfahrungen aus der Praxis der Universität für Angewandte Kunst Wien. *fteval Journal for Research and Technology Policy Evaluation* (49), pp. 38-41. ISSN 1726-6629. <http://repository.fteval.at/491/>

Streicher, J., Polt, W., & Unger, M. (2020). Eine Untersuchung der Marktsituation im Bereich der FTI-Evaluierung in Österreich. *Fteval Journal for Research and Technology Policy Evaluation*, 50, 72–81. <https://doi.org/10.22163/fteval.2020.472>

AUTORINNEN

ISABELLA E. WAGNER

Österreichische Plattform für Forschungs- und Technologiepolitikevaluierung (fteval)

Austrian Platform for Research and Technology Policy Evaluation

Linke Wienzeile 246, 1150 Wien

E: wagner@fteval.at

ORCID: 0000-0002-2772-6771

ELISABETH FROSCHAUER-NEUHAUSER

AQ Austria - Agentur für Qualitätssicherung und Akkreditierung Austria

Bereich Audit & Beratung

Franz-Klein-Gasse 5, 1190 Wien

E: Elisabeth.Froschauer-Neuhauser@aq.ac.at

ORCID: 0000-0002-4586-8706

KATHARINA WARTA

technopolis group | austria

Rudolfsplatz 12/111010 Wien

E: warta@technopolis-group.com

ORCID: 0000-0002-5937-0044

DANKSAGUNG

Die Autorinnen bedanken sich bei den Vortragenden, Discussants, und der Moderatorin – Meike Olbrecht (AvH), Thomas König (IHS), Patrick Lehner (LBG), Michael Hofer (Universität Wien), Brigitte Müller (CDG), Michael Strassnig (WWTF), Elke Dall (ZSI) – und für die Beiträge der Teilnehmenden der fteval-Veranstaltung "Verbindlichkeit in der internen Evaluierung".

SCHLAGWORTE

Interne Evaluierung, Verbindlichkeit, Verantwortung, Beauftragung, Durchführung, Objektivität

KEYWORDS

internal evaluation, commitment, accountability, assignment, execution, objectivity