JOANNEUM RESEARCH

InTeReg Working Paper No. 11-2003

DIE IMPLEMENTIERUNG VON EVALUIERUNGSSYSTEMEN IN FTE PROGRAMMEN

KONZEPTIONELLE UND OPERATIONALE ÜBERLEGUNGEN AN HAND DES ÖSTERREICHISCHEN SPIN OFF PROGRAMMES AplusB

Klaus Zinöcker

InTeReg Working Paper No. 11-2003

DIE IMPLEMENTIERUNG VON EVALUIERUNGSSYSTEMEN IN FTE PROGRAMMEN

KONZEPTIONELLE UND OPERATIONALE ÜBERLEGUNGEN AN HAND DES ÖSTERREICHISCHEN SPIN OFF PROGRAMMES AplusB

Klaus Zinöcker

Joanneum Research, Institute of Technology and Regional Policy
Wiedner Hauptstrasse 76, 1040 Vienna, Austria
e-mail: klaus.zinoecker@joanneum.at
tel: +43-1-581 75 20/2827

Abstract:

Evaluierungssysteme haben zum Ziel, unterschiedliche Evaluierungsschritte auf Programm- und Projektebene zu strukturieren, sie zusammenzuführen, zu optimieren und zu systematisieren und einen Bezug zu Monitoring und Reporting herzustellen. In Evaluierungssystemen werden Gegenstand (Programm, Projekt, Politik), Phasen (ex ante, interim, expost), als auch Inhalte (Konzept-, Design, Prozess- und Wirkungsevaluierung) aufeinander abgestimmt und operationalisiert. An hand des Evaluierungssystems des österreichischen Spin off Projektes AplusB wird die Implementierung eines solchen Systems illustriert.

Keywords: Evaluation, R&D, spin offs, Evaluation Systems

JEL Classification: O32, ,H29, H43

1 Einleitung

„Quis, quid, ubi, quibus auxiliis, cur, quomodo, quando?“[1]

Evaluierungssysteme haben zum Ziel, unterschiedliche Evaluierungsschritte auf Programm- und Projektebene zu strukturieren, sie zusammenzuführen, zu optimieren und zu systematisieren und einen Bezug zu Monitoring und Reporting herzustellen. In Evaluierungssystemen werden Gegenstand (Programm, Projekt, Politik), Phasen (ex ante, interim, expost), als auch Inhalte (Konzept-, Design, Prozess- und Wirkungsevaluierung) aufeinander abgestimmt und operationalisiert.

„Evaluation system: All the rules, institutions, procedures and resources which organize the evaluative function in a given institutional context.[2]

Basis des vorliegenden Papiers war eine Kooperation zwischen der Joanneum Research Forschungsgesellschaft mbH und der Technologie Impulse Gesellschaft mbH zur Erarbeitung von Grundsätzen zu einem Evaluierungssystem von A*plus*B, einem Programm zur Unterstützung akademischer spin offs in Österreich. Im Rahmen dieses Projektes wurden konzeptionelle Überlegungen zum Einsatz von Evaluierung im Projektverlauf von A*plus*B angestellt, es wurde ein Ziel-Indikatorsystem erarbeitet und es wurden Fragebogen für das Monitoring entwickelt. Vorliegendes Working Paper nimmt nun dieses Projekt zum Anlass, um an Hand des Beispiels „Evaluierungssystem A*plus*B“ generelle Überlegungen zu der Implementierung von Evaluierungssystemen in FTE Programmen anzustellen.

Zu Beginn dieses Working Papers wird die Motivation erläutert, die die Implementierung von Evaluierungssystemen in FTE – Programmen begründet. Evaluierungssysteme minimieren Realitätsferne – einerseits von EvaluatorInnen, die vorbei an den Bedürfnissen fordern und produzieren, andererseits der Policy Maker, die wiederum vorbei an den Erfordernissen von Evaluierungen Ergebnisse erwarten. So sind solche Systeme ein wesentliches Mittel, um ·zwischen Programmphilosophie, Implementierung und Programmrealität und Evaluierung zu vermitteln, Best Practice- Überlegungen, Standards der Evaluierung bzw. allgemeine und programm-typenspezifische Evaluierungsrichtlinien bei der Implementierung einer Evaluierung zu berücksichtigen, für alle Beteiligten, vor allem für die Evaluierten, ein größtmögliches Maß an Transparenz zu garantieren sowie den Evaluierungsaufwand zu optimieren und somit eine ‚Evaluierungs- Fatigue’ zu vermeiden.

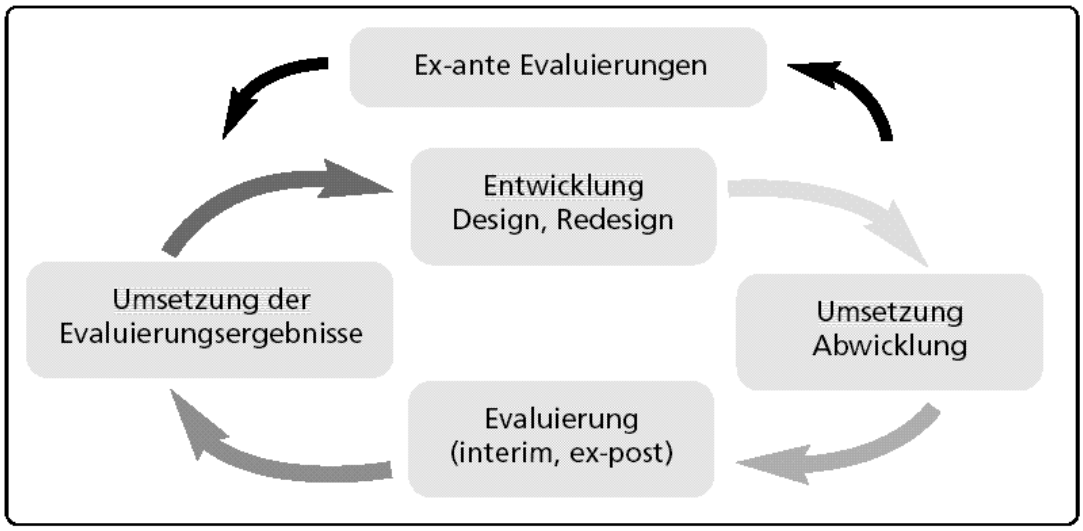
Die Implementierung eines Evaluierungssystems in FTE-Programmen wird hier an einem konkreten Beispiel erläutert - Kapitel 3 illustriert folgerichtig kurz die Funktionsweise dieses Beispiels, dem Spin off Programm der TIG, A*plus*B. Wesentliches Mittel dieser Illustration ist eine Logic Chart Analyse, die sich als zentrales Element an mehreren Stellen dieses Papiers findet.

Evaluierungssysteme ermöglichen die Einbeziehung von konzeptionellen Überlegungen in Organisation und Ablauf einer Evaluierung. Im Entstehungsprozess eines Evaluierungssystems ist Klarheit zu schaffen hinsichtlich Funktion, Zeitpunkten, Phasen und Inhalte der Evaluierung. Ein Profil für jene, die (zu verschiedenen Zeitpunkten) evaluieren, ist zu erstellen und Konsequenzen aus den einzelnen Evaluierungsschritten sind festzulegen. Kapitel 4 hat diese Überlegungen zum Inhalt.

[1] Römischer Zenturio im Hause des Präfekten von Lutetia, Gracchus Überdrus. Réné Goscinny, Albert Uderzo, die Goldene Sichel.
[2] Means Collection Volume 6, Seite 43

Kapitel 5 beschäftigt sich abschließend mit dem eigentlichen Evaluierungssystem A*plus*B. Es beginnt mit einer Darstellung des Monitoring und des Controlling der Zentren, beides die wesentlichen Grundlagen des Systems. Es stellt die fixierten Eckpunkte der Zentrenevaluierung dar und geht detailliert auf das Ziel- Indikatorensystem von A*plus*B ein.

Abbildung 1: FTE-Programme im Politikzyklus



Quelle: Plattform [2003]: Standards der Evaluierung in Forschungs- und Technologiepolitik

Mit diesem Evaluierungssystem wird versucht, einen idealtypischen Politikzyklus für das AplusB Programm nachzuzeichnen und die Vorbereitung zur Umsetzung zu legen: Ein Programm wird auf Grundlage einer Problemstellung entworfen und entwickelt. Die Ex-ante Evaluierung überprüft Angemessenheit und Design des Programms. Nach allfälligen Modifikationen wird das Programm umgesetzt – d.h. in der Regel an eine mit Abwicklung und Programm-Management befasste Einrichtung abgegeben. Eine interim- Evaluierung zieht Zwischenbilanz: Sie hat sowohl den Stand der Projekte als auch das Programm-Management im Blick und gibt Empfehlungen für die Weiterführung. Die Umsetzung der Evaluationsergebnisse – und zwar sowohl der Zwischenevaluierung als auch der Endevaluierung hat Konsequenzen: Das betreffende Programm wird weitergeführt oder auch nicht, es wird modifiziert oder auch nicht, es wird größer oder kleiner gestaltet etc.

Dieses Working Paper versteht sich als Illustration (und vielleicht Anleitung) einer solchen Umsetzung.

Besonderer Dank gilt Dorothea Sturn (TIG) sowie Wolfgang Pointner (Oesterreichsche Nationalbank) für wertvolle Hilfe und Anregungen.

2 Warum Evaluierungssysteme?

Das Matching der Vorstellungen und Intentionen von Policy Maker bzw. AuftraggeberInnen und EvaluatorInnen stößt des öfteren auf beinahe unüberwindbare Probleme. Georghiou[3] bezeichnet das Auseinanderklaffen zwischen den Erwartungen des Auftraggebers und den Auffassungen der EvaluatorInnen als „delivery gap"; die Diskrepanz zwischen den Ansprüchen des EvaluatorInnenteams und den Anforderungen des Auftraggebers als „Customer Gap".

Tabelle 1: Der 'Delivery Gap'

What policymakers want	What evaluators say
• Information in time for spending decision	• Research may take years to have effects
• Clear attribution of effects to investment	• Linear model is a rare case and additionality is complex to assess
• Independent evidence of research excellence	• Peers defend their subject field & international colleagues
• Key indicators to monitor & benchmark	• Crude regime distorts performance & can be manipulated

Quelle: Boden, M. und Stern, E [2002]: User Perspectives. In „RTD Evaluation Tool Box", Seite 13

Diese „Gaps" ließen sich wie folgt karikieren:

Während Policy Maker rechtzeitig Informationen an hand weniger, aussagekräftiger Informationen klar zuordnen wollen können, betonen Evaluatorinnen, das eine klare Zurechenbarkeit eines Effektes eher rar ist und Aussagen über Wirkungen erst nach mehreren Jahren verlässlich treffbar sind. Um klare Aussagen treffen zu können, fordern EvaluatorInnen klare Zielsysteme und eine genügende Ressourcenausstattung, während für Policy Maker eine möglichst rasche Verfügbarkeit von Information von besonderer Notwendigkeit ist.

Tabelle 2: Der 'Customer Gap'

What evaluators want	What policymakers say
• Clearly defined & hierarchical objectives	• Programmes are a compromise involving multiple & conflicting objectives
• Guaranteed independence	• Recommendations must be within realistic policy constraints
• Time & resources to do the job	• We need the results in three months
• Full access to information and stakeholders	• Everyone is overworked and busy

Quelle: Boden, M. und Stern, E [2002]: User Perspectives. In „RTD Evaluation Tool Box", Seite 13

Neben den von Georghiou und Boden & Stern beschrieben Gaps sind auch Diskrepanzen zwischen EvaluatorInnen und Programmmanagement aufzuzeigen: Das, was ein Evaluator / eine EvaluatorIn unter einer sinnvollen und nützlichen Evaluierung versteht kann sich wesentlich von der Auffassung der ProgrammmanagerInnen unterscheiden.

[3] Luke Georghiou [2001] „The Impact and Utility of Evaluation", Vortrag im Rahmen der "Conference on international best practices in evaluation of research in public institutes and universities", Brüssel, 16. 10. 2001

Tabelle 3: Der 'Manager Gap'

What managers say	What evaluators answer
• Look at this nice development • We collect lots of facts about our project	• Where are the social returns? • Not a single number is a useful additionality measure!

Quelle: Sturn & Zinöcker 2003

Evaluierungssysteme helfen nun, diese Gaps abzubauen, indem sie von Anfang an ermöglichen, Erwartungen an Evaluierung klar zu definieren und zu dokumentieren. Diese Offenlegung der Erwartungen ist für alle stakeholder, also für all jene, die an der Evaluierung berechtigte Interessen haben, von Bedeutung.

2.1 MÖGLICHKEITEN VON EVALUIERUNGSSYSTEMEN

Evaluierungssysteme sind nun ein wesentliches Mittel, um

- zwischen Strategie, Programm, Maßnahmenalltag und Evaluierungsumsetzung zu vermitteln,
- Best Practice- Überlegungen, Standards der Evaluierung bzw. allgemeine und programmtypenspezifische Evaluierungsrichtlinien bei der Implementierung einer Evaluierung zu berücksichtigen,
- für alle Beteiligten ein größtmögliches Maß an Transparenz zu garantieren sowie
- den Evaluierungsaufwand zu optimieren und somit eine ‚Evaluierungs- Fatigue' zu vermeiden.

Vermittlung zwischen Programmrealität und Evaluierungsumsetzung

Ein klarer Kritikpunkt, zwar aus verschiedenen Blickwinkel betrachtet, doch auf einen Nenner gebracht, der sich aus den oben zitierten „Gaps" herauskristallisieren lassen, ist „Realitätsferne": einerseits von EvaluatorInnen, die vorbei an den Bedürfnissen fordern und produzieren, andererseits der Policy Maker, die wiederum vorbei an den Erfordernissen von Evaluierung Ergebnisse erwarten. Ein Matching ist zu einem späten Zeitpunkt nur mehr schwer möglich, daher versuche Evaluierungssysteme, die zu Programmstart implementiert werden, diese Vermittlungslücke zu schließen. Die Entstehung des Evaluierungssystems ist hier als interaktiver Prozess zwischen Policy Maker bzw. Programmmanager einerseits und EvaluierungsexpertInnen andererseits zu verstehen, in dem die Erwartungen an die Evaluierung mit den Umsetzungserfordernissen der Evaluierung abgeglichen werden. Ein Beispiel: Der/die Policy Maker beschließt die Richtlinien eines Programms mit einer beschränkten Laufzeit von 4 Jahren mit Verlängerungsoption. Für eine solche Verlängerungsentscheidung wäre eine Interimsevaluierung ein wichtiges Entscheidungsinstrument. In einem Evaluierungssystem ist nun dafür Sorge zu tragen, dass zu diesem Zeitpunkt eine Interimsevaluierung vorgesehen wird; diese sollte auf Synergien mit den Projektevaluierungen geprüft und (im Aufwand) optimiert werden. EvaluierungsexptertInnen haben im oben beschriebenen interaktiven Prozess nun den Policy Maker auf diese Notwendigkeit hinzuweisen und mögliche Synergieeffekte aufzuzeigen. Ein weiteres Beispiel: Im Zielsystem eines Programms sind sowohl „*mehr Kooperationen*" als auch „*mehr Forschungsprojekte mit hohen Additionalitätseffekten*" verankert. Hierzu weisen die EvaluierungsexpertInnen auf verschieden methodische Ansätze hin, etwa Kontrollgruppenansätze und Netzwerkanalyse, die beide mit hohem (Evaluierungs-)Aufwand umzusetzen wären. Im Diskurs mit Policy Maker und Programmmanagement

werden die „Pros“ und „Cons“ dieser Ansätze an Hand der gewünschten Funktionen der Evaluierung gespiegelt und vor dem Hintergrund, dass der Aufwand für beide Methoden zu hoch wäre, die Vorbereitungen zur Datengenerierung für die gewählte Methode getroffen.[4]

Standards berücksichtigen

Die Implementierung eines Evaluierungssystems zu Beginn eines Programms ermöglichen es, Best-Practice- Arbeiten, allgemeine und programmtypenspezifische Evaluierungsrichtlinien bei der Implementierung einer Evaluierung zu berücksichtigen. Beispiele für solche Arbeiten sind die Standards der Evaluierung von FTE-Programmen der Plattform Forschungs- und Technologieevaluierung GesbR, Arbeiten zur Best Practice Evaluierung von FTE Programmen in Österreich (Fritz et al. 1997) und Papiere wie der RTD Evaluation Toolbox (Fahrenkrog, Polt, Rojo, Tübke, Zinöcker 2002)

Dadurch ist ein mehr an Qualität im Evaluierungssystem und eine Anpassung an ein (europäisches) Niveau der Evaluierung sicherzustellen.

Transparenz

Wonach und von wem der Projekterfolg beurteilt wird ist eine essentielle Information für jedeN ProgrammteilnehmerIn. Bereits zu Beginn eines Programms ist es notwendig, transparent zu zeigen, anhand welcher Kriterien und Indikatoren zu welchem Zeitpunkt evaluiert wird. Damit kann zum einen vermieden werden, dass unterschiedliche Erwartungen und Vorstellungen hinsichtlich der Evaluierung entstehen, zum anderen soll bereits von Anfang an sicher gestellt sein, dass die jeweils notwendigen Daten in adäquater Weise erfasst werden.

Vermeiden von Evaluierungs-Fatigue

Unter FörderempfängerInnen, ProgrammteilnehmerInnen und auch –managerInnen ist ein Klagen über die Beanspruchung im day-to-day business mit (im weiteren Sinne) Informationsaufbereitung für EvaluatorInnen bzw. für Fördervergabestellen oft gehört; vor allem das Ausfüllen von Fragebögen dürfte zu den weniger beliebten Aufgaben in diesem Personenkreis gehören.

Dies kann weder im Sinne des Policy Maker, des Programmmanager noch im Sinne der EvaluatorInnen liegen. So sinkt mit einer zunehmenden Evaluierungs-Müdigkeit nicht nur die Rücklaufquote, sondern auch die Qualität der zur Verfügung gestellten Information. Um dies zu vermeiden ist eine strukturierte Gestaltung des Berichtswesens und des Monitorings eines Programms wesentlich und sollte Teil eines umfassenden Evaluierungssystems sein. Ein Zusammenführen des Informationsbedarfs von Programmmanagement und Evaluation, ein Zusammenführen der Erhebungsschritte vermeidet etwaige Doppelgleisigkeiten und stellt eine wesentliche Arbeitserleichterung für die Beantwortenden dar.[5]

2.2 EVALUIERUNGSSYSTEME ALS INTERAKTIVER PROZESS

Der Entstehungsprozess von Evaluierungssystemen ist als interaktiver Prozess zwischen Policy Maker, Programmmanager und begleitenden EvaluierungsexpertInnen anzulegen. Evaluierung ist kein starrer, eingleisiger Prozess. Abhängig vom Lern- und Lenkungswillen der Beteiligten, abhängig von der Programmart, den Bedürfnissen der beteiligten Personen und den Anforderungen Dritter (in Österreich

[4] vgl. zu diesem Beispiel auch die Anmerkungen in Abschnitt 2.2

etwa des Rates für Forschungs- und Technologiepolitik) sind die Ansprüche an den Einsatz von Evaluierung von Programm zu Programm verschieden. Ohne den daraus resultierenden Bedarf genau zu kennen ist die Entwicklung eines Evaluierungssystems „vergebene Liebesmüh'". Schließlich ist das wesentliche Asset dieser Entwicklung die Berücksichtigung der Programmrealität, und das funktioniert nur im Dialog mit Programmmanager und forschungs- und technologiepolitischen Entscheidungsträgern.

2.3 GRENZEN VON EVALUIERUNGSSYSTEMEN

Unabhängig davon, aus welcher Position heraus der Anstoß zur Schaffung von Evaluierungssystemen erfolgt, muss im Sinne einer „Evaluierungssubsidiarität" das Verständnis vorausgesetzt werden, Rollen nicht auszureizen und Zuständigkeiten nicht überschreiten zu dürfen.

Die Schaffung von Evaluierungssystemen kann sowohl von Policy Maker als auch von Programmmanager und –abwickler ausgehen. Je nach Position des Initiators ist auch die Ausrichtung und der Inhalt des Systems verschieden. Policy Maker, sollten sie Evaluierung als Instrument ernst nehmen, können, ja müssen Evaluierungsschritte des Programms planen, abhängig von der Laufzeit des Programms oder der Programmrichtlinien, müssen Konsequenzen dieser Schritte und Erfolgsindikatoren für das Programm vorgeben. Programmmanager wiederum setzen Evaluierung als Lenkungs- und Informationsinstrument ein.

Was wäre eine Überschreitung der den Policy Makern bzw. den Programmmanagern zukommenden Rollen? Etwa, wenn der Programmmanager die Terms of References für die Evaluierung des Programms gestaltet oder der Policy Maker direkt in die Auswahl und die (Zwischen-)evaluierung der Projekte eingreift. Somit werden auch die Mindesterfordernisse an ein Evaluierungssystem von der Position des Initiators bestimmt.

Eine Beschneidung der Rolle des Evaluators?

Evaluierungssysteme ersetzen in keiner Weise eine eigentliche Programm- oder Projektevaluierung, sie bringen sie nur in eine Ordnung (auch zueinander). In Evaluierungssystemen setzt man sich aber sehr wohl mit Indikatoren zur Bewertung von Erfolg und Misserfolg, mit Konsequenzen von Evaluierung und auch mit Methoden bzw. mit den Datenerfordernissen zu diesen Methoden auseinander. Wenig Wohlmeinende werden hierin eine Beschneidung der Rolle und der Befugnisse des eigentlichen Evaluators, sei es auf Projekt- oder auf Programmebene, sehen. Dies ist an dieser Stelle zu relativieren. In der Evaluierung von Programmen und Projekten ist ein trade off zwischen der - absoluten - Unabhängigkeit des Evaluierungsteams und der Transparenz (der Kriterien zur Erfolgsbemessung) kaum abwendbar; auch ein trade off zwischen Unabhängigkeit und dem Einsatz von Evaluierung als Lenkungsinstrument für Policy Maker und Programmmanager ist feststellbar. Jedenfalls sollte es Aufgabe von Evaluierungssystemen sein, den späteren EvaluatorInnen hinreichende Freiheiten in der Umsetzung der Evaluierung zu gewähren-

Auch ohne der Entwicklung von Evaluierungssystemen ist dieser trade off zu verzeichnen: Terms of References geben EvaluatorInnen detaillierte Aufgabenstellungen[6] und schließen a priori mögliche Fragestellungen aus.

---

[5] "Evaluation fatigue" can be a problem with project participants who have to respond to many different questionnaires, especially when these ask repeatedly for the same data. Evaluation fatigue could be avoided by co-ordinating the gathering of data for monitoring and evaluation studies with the information needs of other RTD policymaking bodies." ETAN Working Paper 1999

[6] Oft wird eine Konzentration auf gewisse Evaluierungsphasen – z.B. Prozessevaluierung – oder explizit ein methodischer Ansatz vorgegeben.

3 Academia plus Business – A*plus*B

Nachfolgendes Kapitel soll kurz die Funktionsweise des Academia plus Business Programms A*plus*B (www.aplusb.at) der Technologie Impulse Gesellschaft TIG (www.tig.or.at) illustrieren. AplusB dient schließlich als Illustration für das vorliegende Papier, so soll ein gewisses Grundverständnis für das Programm hinter der Evaluierung erzeugt werden.

Zur Motivation eines Spin off Programms für Österreich

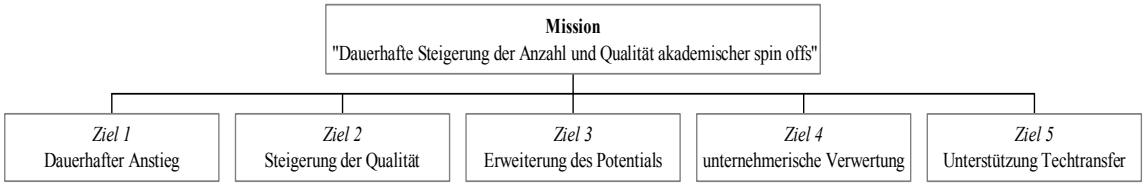
Österreich weist im internationalen Vergleich eine eher mäßige Gründungsdynamik auf. Besondere Schwächen zeigen sich dabei in höheren Technologiesegmenten: Weniger als 10% der Unternehmensgründungen kommen aus dem "High Tech Bereich" Die Zahl der GründerInnen mit abgeschlossenem Universitätsstudium und wissenschaftlichem Hintergrund ist signifikant niedrig. [7] Das Impulsprogramm A*plus*B verfolgt daher das Ziel, die Zahl der innovativen und technologieorientierten Unternehmensgründungen aus dem akademischen Sektor dauerhaft zu steigern.

Ziele von AplusB

Die Ziele des AplusB Programms sind in den Richtlinien des Programmes festgelegt:

> Ziel [der Förderaktion] ist die dauerhafte Steigerung der Zahl und des Erfolgs von Gründungen aus Universitäten, Fachhochschulen und außeruniversitären Forschungseinrichtungen. Daraus ergeben sich die folgenden Teilziele:
>
> - Dauerhafter Anstieg der Anzahl akademischer Spin-offs
> - Steigerung der Qualität (Technologie- und Wissensintensität) und Erfolgswahrscheinlichkeit dieser Gründungen
> - Erweiterung des Potentials an Unternehmensgründungen aus Universitäten, Fachhochschulen und außeruniversitären Forschungseinrichtungen.
> - Verbesserung der unternehmerischen Verwertung von Forschungsergebnissen
> - Unterstützung anderer Maßnahmen des Technologietransfers.

Abbildung 2: Programmziele A plus B



Die Zentren selbst müssen darüber hinaus Ziele und Zielgruppen definieren, die dem Zentrum ein kompaktes, konsistentes, unverwechselbares Aufgabenprofil geben. Zu berücksichtigen sind hierbei das Profil der Partner, die regionale Situation und das regionale Potential, und bereits vorhandene Ansätze und Institutionen der Gründerförderung, des Technologietransfers und eine bestmögliche Nutzung regionaler Synergien.

[7] Vgl. dazu Sturn [1999]

Zentrenziele sind hierbei Vertiefung oder Konkretisierung der Programmziele, aber auch eine Erweiterung (wenn man so will, Ergänzung). Die Zentren verpflichten sich durch den Abschluss einer Fördervereinbarung den Richtlinien des A*plus*B Programms, und somit den darin enthaltenen Programmzielen. Der Erfolg eines Zentrums, der Grad seiner Zielerreichung bemisst sich also an der Summe von Zentren- und Programmzielen.

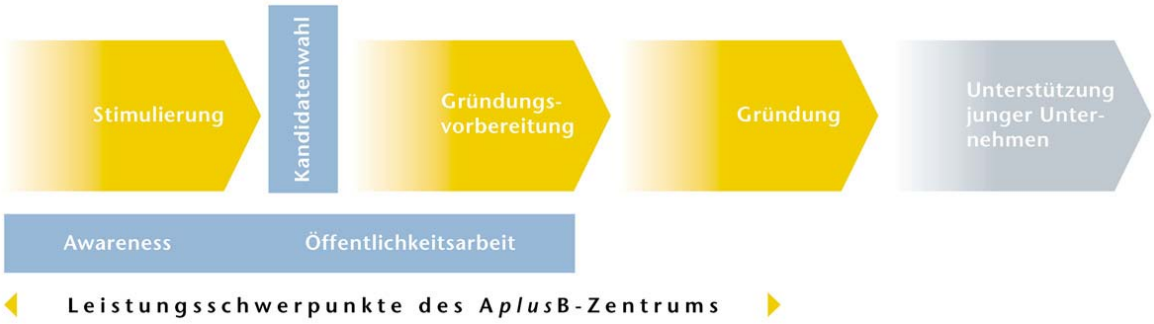
A plus B-ZENTREN

A*plus*B-Zentren unterstützen die Vorbereitung von Gründungen aus Universitäten, Fachhochschulen und Forschungseinrichtungen. WissenschaftlerInnen bekommen hier die Möglichkeit, sich auf dem schwierigen Weg von einer guten Idee zu einem Unternehmen professionell begleiten zu lassen. Dabei geht es zum einen um konkrete Beratung und Unterstützung im Vorgründungsprozess, zum zweiten aber auch darum, Unternehmertum als Option im akademischen Denken und Handeln stärker zu verankern. Die enge Anbindung der GründerInnen an ihre wissenschaftliche Heimat stellt sicher, dass die neu geschaffenen Unternehmen jenes Know-how wirtschaftlich umsetzen, welches an akademischen Einrichtungen entwickelt wird. Auf diese Weise leisten A plus B-Zentren auch einen wichtigen Beitrag zur Intensivierung und Verfestigung der Forschungskooperationen zwischen Wissenschaft und Wirtschaft.

AKTIVITÄTEN VON A plus B ZENTREN

A *plus* B-Zentren konzipieren ein aufeinander abgestimmtes Bündel an Maßnahmen für junge WissenschaftlerInnen: Mobilisierung und Stimulierung von Gründungen, Interesse wecken (Veranstaltungen, Informationsarbeit, Lehrstühle für Entrepreneurship etc..), Beratung, Qualifikation und Betreuung für 1,5 Jahre (wissenschaftlich-fachliche Betreuung und Coaching, Management-Beratung, Weiterbildung) sowie eine Optimierung der Startbedingungen für die jungen Unternehmen durch entsprechende Kooperationen mit Financiers und mit anderen Programmen.

Abbildung 3: Leistungsspektrum des A*plus*B *Zentrums*



Quelle AplusB Academia plus Businness Präsentation, TIG, September 2001

3.1 VERSUCH EINER LOGIC CHART ANALYSE

Logic Charts sind Diagramme, die die Beziehungen zwischen Zielen, Aktivitäten, outputs, outcomes, impacts[8] eines Programms darstellen. Sie visualisieren das Design, die zugrundeliegenden Annahmen und verbinden es mit gewünschten Effekten. Vorteil von solchen Diagrammen ist die Möglichkeit zur

[8] Vgl. dazu Box 1: Outputs, outcomes, impact weiter unten.

intuitiven Erfassung einer Programmlogik. Wenn allerdings versucht wird, komplexe Beziehungsgeflechte möglichst wahrheitsgetreu abzubilden, so reizt man schnell den Charme dieses Visualisierungsansatzes aus.

Box 1: Outputs, outcomes, impact

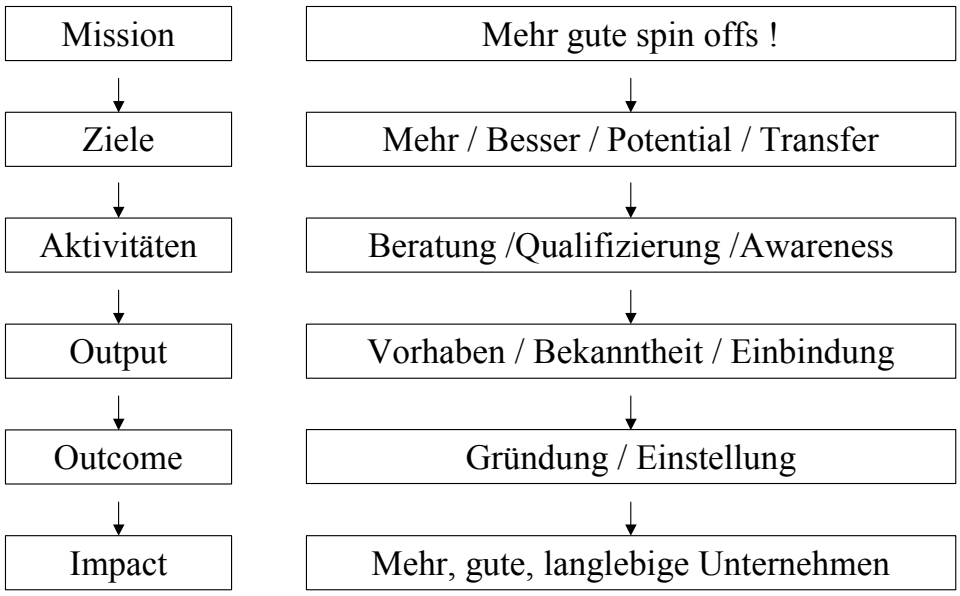
Outputs:	the technical results of the projects
	Im Falle von AplusB: Gründungsvorhaben, Stimulierungsaktivitäten
Outcomes:	the direct effects of the project
	Im Falle von AplusB: Gründungen
Impact:	the wider effects of the programme on the society.
	Im Falle von AplusB: Dauerhafter Anstieg der Gründungen in Österreich

In principle, outputs cause outcomes, and outcomes cause impacts.

Quelle: nach Erik Arnold und Ken Guy: Technology Diffusion Programms and the Challenge for Evaluation. IN. OECD 1997, Seite 81

In Abbildung 4 wird daher eine solche Logic Chart für A*plus*B stark vereinfachend und auf „Slogans" reduzierend dargestellt.

Abbildung 4: Schematische Darstellung einer Logic Chart für AplusB



Obige Abbildung stellt nun den Kontext für weitere Überlegungen zum Evaluierungssystem von A*plus*B in den folgenden Kapiteln dar. Die Logic Chart Analyse wird im Abschnitt 5.4 weitergeführt.

4 Konzeptionelle Überlegungen zum Evaluierungssystem von A*plus*B

Evaluierungssysteme ermöglichen die Einbeziehung von konzeptionellen Überlegungen in Organisation und Ablauf einer Evaluierung. Im Entstehungsprozess eines Evaluierungssystems ist Klarheit zu schaffen hinsichtlich Funktion, Zeitpunkten, Phasen und Inhalte der Evaluierung. Ein Profil für jene, die (zu verschiedenen Zeitpunkten) evaluieren, ist zu erstellen und Konsequenzen aus den einzelnen Evaluierungsschritten sind festzulegen.

Um den Entstehungsprozess des Evaluierungssystems A*plus*B besser illustrieren zu können ist es vorerst notwendig, ein klares Bild der Rolle der Beteiligten zu entwerfen.

Im Jahr 2001 wurde das Management des A*plus*B Programms der TIG durch das Bundesministerium für Verkehr, Innovation und Technologie übertragen. Die TIG um initiiert Wettbewerbe, bei denen sich unterschiedliche Konsortien um eine Förderung bewerben können. Eine unabhängige ExpertInnenjury bewertet die Qualität der eingereichten Vorhaben und formuliert eine Förderungsempfehlung an die/den zuständigen BundesministerIn. Zentren fördern nun Gründungsvorhaben, betreuen diese für einen gewissen Zeitraum und unterstützen diese auch nach der erfolgten Gründung.

Hieraus ergeben sich nun für ein Evaluierungssystem verschiedene Rollenzuschreibungen und damit Sprachregelungen:

Policy Maker	ist im Rahmen des AplusB Programms das zuständige Ministerium, das bmvit[9].
Programmmanager	des AplusB Programms ist die TIG.
Programm	ist klarerweise *Aplus*B.
Projekte	Sind die einzelnen Zentren, die im Rahmen von AplusB gefördert werden.

Auf Projekt- bzw. auf Zentrenebene sind vier Evaluierungsschritte geplant bzw. in den Richtlinien vorgegeben: Die Gestaltung dieser Evaluierungen liegt in der Verantwortung des Programmmanagers, der TIG. Ob das Programm zwischen- und endevaluiert wird liegt in der Verantwortung des Policy Makers, dem bmvit, wobei Anregungen von Seiten des Programmmanagers möglich sind.

Diese Unterscheidung und klare Trennung der Funktionen ist für die Entwicklung des Evaluierungssystems von Bedeutung, gibt sie doch die Schwerpunktsetzung vor: Das Evaluierungssystem hat sich auf die Projektebene zu konzentrieren und für Evaluierungsschritte auf Programmebene zum jetzigen Zeitpunkt möglichst weitreichende Vorarbeiten zu leisten.

Für die konzeptionellen Überlegungen zum Evaluierungssystem hat diese Schwerpunktsetzung vorerst keine Bedeutung. (Gute) Erfahrungen aus dem Bereich der Evaluierung arbeitsmarktpolitischer Projekte[10] haben gezeigt, dass „Good“ oder „Best Practice“ Überlegungen im Bereich der Evaluierung gleichsam für Projekt- als auch auf die Programmebene angewandt werden. Daher wird nachfolgend versucht, solche "Best Practice - Überlegungen" für FTI - Programme zu nutzen und zu untersuchen, welche Schlussfolgerungen daraus für das Evaluierungssystem in A *plus* B zu ziehen ist.

[9] Mit Abstrichen auch verschiedene Landesstellen, welche die in ihren Bundesländern gelegenen Zentren kofinanzieren.
[10] Schmid [1997]

Zeitliche Dimension der Evaluierung

Eine gebräuchliche Einteilung von Evaluierungen ist jene in ex ante, interim und expost und setzt am Zeitpunkt der Durchführung in Bezug zum Programmverlauf an:

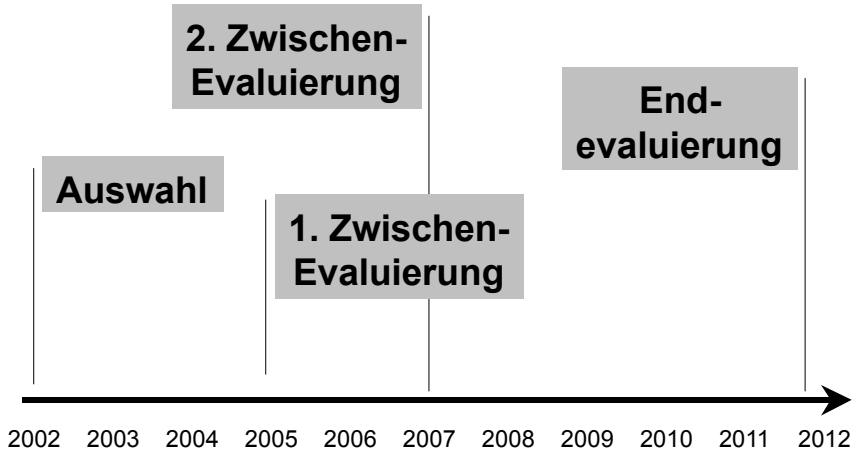
- *Ex-ante* Evaluierungen setzen vor Programmbeginn an, haben einen zukunftsgerichteten Charakter und ermöglichen eine erweiterte „rationality in programmes".
- *Interim* Evaluierungen werden während der Laufzeit eines Programms durchgeführt.
- *Ex-post* Evaluierungen starten nach Programmende und betrachten die Programmentwicklung, Impacts und Kosten bzw. Nutzen (sofern möglich) retrospektiv.

Bereits in den Richtlinien des A*plus*B Programms wird detailliert auf die Projekt- bzw. Zentrenevaluierung eingegangen. Es ist von vier Auswahl-, Bewertungs-, bzw. Evaluierungszeitpunkten auf Projekt- bzw. Zentrenebene auszugehen, die in den Richtlinien festgelegt worden sind:

1. Bewertung der Förderansuchen im Rahmen eines Auswahl-Wettbewerbes
2. erste Zwischenevaluierung im dritten Jahr
3. zweite Zwischenevaluierung im fünften Jahr
4. Abschließende Evaluierung nach zehn Jahren.

Für das erste Auswahlverfahren (abgeschlossen 2002) hieße dies, dass mit folgenden Evaluierungszeitpunkten zu rechnen ist:

Abbildung 5: Evaluierungszeitpunkte in AplusB auf Projektebene



Zeitpunkt für die erste Zwischenevaluierung wäre also 2005, die zweite (Stop or Go) folgte 2007, bei einer Go-Entscheidung würde die Endevaluierung der Zentren 2012 erfolgen. Sollte nun angedacht werden, auch das Programm zwischen zu evaluieren, wäre demnach der ehestmöglich geeignete Zeitpunkt 2006.

Fügte man Programmevaluierungsschritte in *AplusB* in dieses System ein, könnte man von einem „rollierenden" System der Evaluierung (mit einer hohen Lern- und Informationskomponente sowohl für Programmmanager als auch Policy Maker) ausgehen: Das Auswahlverfahren bewertet die einzelnen Zentren nach den in den Richtlinien festgelegten Kriterien. („ex ante") In der ersten Zwischenevaluierung („interim") wird die Arbeit des Zentrums bewertet und Empfehlungen für die Zukunft ausgesprochen. Die 2. Zwischenevaluierung trifft eine Stop and Go Entscheidung („ex post") und bewertet

gleichzeitig die Pläne für die Jahre 5 – 10 (wiederum „ex ante"). Nach 10 Jahren werden die Zentren abschließend bewertet. („ex post").

Tabelle 4: Ein rollierendes Evaluierungsmodell in AplusB

	Auswahl	1. Zwischen- Evaluierung	2. Zwischen- Evaluierung	End-Evaluierung
2002	ex ante			
2005		interim		
2007			ex post - ex ante	
2012				Ex post
Programmevaluierung				

Funktion der Evaluierung

Die Evaluation öffentlicher Programme kann, abhängig von den jeweiligen Rahmenbedingungen, verschiedene Funktionen erfüllen:

- eine *Legitimationfunktion* – eine Rechtfertigung der Verwendung von öffentlichen Mittel.
- eine *Informationsfunktion,* - eine Information an die Öffentlichkeit über den Umgang mit und die Verwendung von öffentlichen Mittel,
- eine (informationsorientierte) *Lernfunktion*, für Programmträger, Träger technologiepolitischer Entscheidungen, WissenschafterInnen etc-
- eine *Lenkungsfunktion* für die zukünftige Gestaltung von Politikzielen, von Maßnahmenplanung etc.,
- und eine *Kontrollfunktion* im Sinne eines Controllings eines privatwirtschaftlichen Unternehmens.

Nicht zu jedem Zeitpunkt muss oder soll Evaluierung alle hier skizzierten Funktionen ausfüllen können. Aufgabe eines Evaluierungssystems ist es auch zu bestimmen, zu welchem Zeitpunkt Evaluierung welche Funktion auszufüllen hat.

Die geplanten Evaluierungsschritte in A*plus*B sind, wie weiter oben in Abbildung 5 bzw. in Tabelle 4 skizziert, im Evaluierungssystem auf Ebene der Projekte bzw. der Zentren geplant. Fragen nach der *Legitimation* des Einsatzes öffentlicher Mittel für die Förderung eines Spin-off Gründerprogramms sind zum Teil in der Feasibility Studie zu A*plus*B beantwortet und können mit Evaluierungen auf Projekt- bzw. auf Zentrenebene (noch) nicht in der Beantwortung ergänzt werden, genauso können aus ihnen keine unmittelbaren Konsequenzen für die zukünftige Politikgestaltung gezogen werden. Für eine Legitimationsfunktion benötigte es zusätzlich einer Programmevaluierung.

Die Evaluierungsschritte auf Zentrumsebene sind in den Richtlinien zu verschiedenen Zeitpunkten (und mit unterschiedlichen Konsequenzen) geplant, was dem Programmmanagement die Möglichkeit zur Umsetzung der Evaluierungsergebnisse gibt, also zu verschiedenen Zeitpunkten *lenkend* und sicher auch *lernend* in die Arbeit des Zentrums einzugreifen.

Das geplante Berichtswesen (vgl. dazu Abschnitt x weiter unten) bietet ausreichende *Information* und *Kontroll*möglichkeiten sowohl für Policy Maker als auf für das Programmmanagement.

Inhalte und Phasen der Evaluierung

Die Konstruktion eines idealtypischen inhaltlichen Rahmens in Best Practice Ansätzen einer Evaluation[11] bringt eine Unterscheidung in Konzept-, Design-, Prozess-, und Wirkungsevaluierung mit sich.

Konzeptevaluierungen überprüfen Mission, Annahmen, grundlegende Thesen und Rahmenbedingungen der Programme (Institutionen, Projekte). Bewertet wird hierbei, wie berechtigt der Einsatz der Politik, des Programms/Projektes (inklusive der darin enthaltenen Lösungsansätze) zur Bekämpfung der vorher identifizierten Probleme ist. Konzeptevaluierungen sind typischerweise Elemente von Ex-ante Evaluierungen.

Designevaluierungen beschäftigen sich mit der Zweckmäßigkeit der Ausgestaltung eines Programms und seiner organisatorischen Struktur. Dabei wird überprüft, inwieweit die vorgesehenen operationalen Ziele den jeweiligen Problemsituationen entsprechen, ob die eingesetzten Instrumente diesen Zielsetzungen auch angemessen sind und die angestrebten Ziele auch erreicht werden können bzw. konnten. Darüber hinaus soll auch die organisatorische Struktur, also alle geplanten Aktionen sowie deren Abfolge hinsichtlich ihrer zeitlichen Konsistenz sowie der Übereinstimmung mit dem gewählten Lösungsansatz bewertet werden. Designevaluierungen sind typischerweise Elemente der Evaluierung von Programmen zu allen Zeitpunkten (Ex-ante, Interim und Ex-post).

Das Abrücken der Evaluierung von einem blinden Gegenüberstellen der Situation vor und nach einer Maßnahme („black box Evaluierung“) ermöglicht es, im Rahmen der *Prozessevaluierung* die Durchführung v. a. eines Programms im Detail zu untersuchen. Themen sind hier die Qualität der Implementierung, die laufende Kontrolle des Erfolgs sowie Effizienz und Effektivität der Abläufe. Prozessevaluierungen sind typischerweise Elemente von Interim-Evaluierungen und stehen bei Programmen mit hohen Anforderungen an das Programm- und Projektmanagement im Vordergrund.

Wirkungsanalysen schließlich bewerten die Erreichung der Programmziele (Effektivitätsanalyse) und versuchen alle durch das Programm direkt oder indirekt ausgelösten Effekte zu identifizieren und möglichst auch in ökonomischen Einheiten zu quantifizieren (z. B. induzierte Umsatzsteigerungen etc.). Im Zentrum steht die Erfassung der sozialen, über die privaten hinausgehenden, Erträge des Programms/Projekts. Wirkungsanalysen sind wichtige Elemente von Ex-post Evaluierungen.

Die tatsächlich behandelten Inhalte hängen vom Programmtyp, den Funktionen, den zur Verfügung stehenden Methoden und nicht zuletzt von den Auftraggebern, kurzum von der Programmrealität ab. Klar ist, dass Konzeptevaluierung Sache von Programmevaluierung, nicht der Projektevaluierung ist: Fragen nach der Interventionslegitimation, danach, ob politische und wirtschaftliche Rahmenbedingungen genügend berücksichtigt sind oder Motivationsanalysen sind projektübergreifend und somit einer Programmevaluierung zuzuordnen. Im Rahmen des Evaluierungssystems für A*plus*B heißt das, dass alle entsprechenden Fragestellungen in den anstehenden Evaluierungen auf Zentrenebene auszuklammern sind, da ansonsten eine unsinnige Duplizität mit Aufgaben der Programmevaluierung vorliegt. Gleichzeitig sind in den Feasibility – Studien zum Programm A*plus*B diese – aus einem ex ante Aspekt heraus – abgedeckt.[12]

Designevaluierung ist Sache der Evaluierungsschritte nach drei bzw. fünf Jahren auf Zentrumsebene. Da die A *plus* B-Zentren selbst ein aufeinander abgestimmtes Bündel an Maßnahmen konzipieren, also eine hohe Gestaltungsfreiheit seitens der Zentren besteht, ist ein Hinterfragen der Zweckmäßigkeit und Adäquatheit dieses Bündels zu diesen Zeitpunkten unbedingt notwendig. Gleiches gilt für die Prozessevaluierung.

[11] Fritz et al. 1997

Für die Prozessevaluierung ist es allerdings notwendig, zu einer Abgrenzung zur Tätigkeit des Controllings der TIG zu kommen. Überschneidungen, wie sie in der laufenden Kontrolle des Erfolgs sowie Effizienz und Effektivität der Abläufe gut möglich sind, sind zu vermeiden.

Es wird notwendig sein, eine Abstufung der Inhalte der Evaluierung zwischen den Schritten nach drei und nach fünf Jahren zu finden. Während es notwendig sein wird, nach drei Jahren sämtliche mögliche Fragestellungen aus Design- und Prozessevaluierung zu beantworten, könnte man sich im Rahmen des Evaluierungsschrittes nach fünf Jahren nur mehr darauf konzentrieren zu überprüfen, ob in Folge etwaig identifizierter Missstände Adaptionen des Designs vorgenommen wurden, ob diese zielführend u. zweckmäßig waren oder ob Empfehlungen im Rahmen der Prozessevaluierung umgesetzt worden sind.

Wirkungsevaluation ist sowohl Teil einer Programmevaluation als auch Teil der Zentrenevaluation. Um in einer Evaluierung die Wirkungen eines Programms beurteilen zu können, werden üblicherweise und vor allem die Veränderungen bei den geförderten Einheiten (im Falle AplusB wären dies die Unternehmensgründungen) im Unterschied zu einem (hypothetischen) Zustand ohne Förderung erhoben. Damit sollen jene Effekte gemessen werden, die sich kausal auf die Programmteilnahme zurückführen lassen, die also nicht eingetreten wären, wenn die betreffenden Einheiten nicht an dem Programm teilgenommen hätten (Additionalität). Abhängig von der zur Verfügung stehenden Datengrundlage lassen sich die Wirkungen eines Programms in unterschiedlichen Spezifikationsgraden messen. Um die Wirkungen eines Programms beurteilen zu können, müssen jene Veränderungen der geförderten Einheiten erhoben werden, die sich in einem kausalen Zusammenhang von der Programmteilnahme ableiten lassen. Diese Änderungen von Merkmalen der geförderten Einheiten oder Indikatoren im Zeitablauf werden durch einen Vorher-Nachher Vergleich identifiziert, also durch den Unterschied in diesen Variablen vor und nach der Programmteilnahme.

Dieser Vorher-Nachher Vergleich sagt allerdings noch nichts über die Effektivität einer untersuchten Einheit aus, da die Änderungen der betrachteten Merkmale nicht notwendigerweise auf den Einfluss des Programms zurückzuführen sind, sondern beispielsweise auch auf einen gesamtwirtschaftlichen Trend, der nicht geförderte Einheiten in gleichem Ausmaß erfasst haben kann. Um die Zurechenbarkeit der durch ein Programm induzierten Änderungen von solchen allgemein auftretenden zu gewährleisten, wird ein Kontrollgruppenansatz verwendet. Dabei werden die Änderungen der geförderten Einheiten denen von nicht geförderten gegen-übergestellt. Der Effekt der Förderung ergibt sich aus der Differenz der gemessenen Änderungen.

Durch den Kontrollgruppenansatz können zwar unterschiedliche Änderungen bei TeilnehmerInnen und Nicht-TeilnehmerInnen festgestellt werden, nicht aber eine kausale Beziehung zwischen dem Programm und den aufgetretenen Änderungen. Es besteht noch immer die Möglichkeit, dass die geförderten Einheiten auch ohne Programm Aktivitäten gesetzt hätten, die zu den beobachteten Effekten geführt hätten. In diesem Fall spricht man von einem "selection bias", d.h. in einer Evaluierung wird nicht nur die Wirkung eines Programms bewertet, sondern zum Teil auch das, was die geförderten Einheiten ohne Förderung umgesetzt hätten. Im Bereich des A*plus*B Programms könnte man z.B. annehmen, dass Personen mit hoher Affinität gegenüber selbstständigem Arbeiten auch über eine überdurchschnittliche Kenntnis der Gründungsförderlandschaft verfügen, etwa weil sie ohnehin fertige Gründungskonzepte vorliegen haben. Die Unterstützung von Personen, die auch ohne Förderung ein Unternehmen professionell gegründet hätten, produziert ein hohes Ausmaß an Mitnahmeeffekten und verringert daher die Additionalität des Programms.

---

[12] Sturn [1999]

Um auch den "selection bias" bei einer Evaluierung berücksichtigen zu können, werden mikroökonometrische Methoden eingesetzt. Diese Methoden machen Gebrauch von Zusatzinformationen über geförderte wie nicht geförderte Einheiten, die einen Einfluss auf die Teilnahme an einem Programm haben. Dadurch lassen sich Effekte isolieren, die nur auf den Einfluss des Programms zurückzuführen sind und nicht auch auf die Auswahl der geförderten Einheiten. Das Ergebnis dieses Ansatzes gibt Auskunft über die tatsächlich erzielte Additionalität, ohne die bei einem herkömmlichen Kontrollgruppenansatz unvermeidliche Verzerrung. Additionalität besteht aus der Differenz zwischen der Entwicklung, die eine geförderte Einheit nimmt, zu der Entwicklung, die diese Einheit ohne Förderung genommen hätte. Mit mikroökonometrischen Methoden kann ein solcher Referenzwert simuliert werden.

Dazu ist es aber unerlässlich, dass bereits vor dem Zeitpunkt der Ex Post Evaluierung entsprechende Daten erhoben werden, damit eine Veränderung im Zeitablauf gemessen werden kann. Wenn Daten über beide Zeitpunkte (vorher und nachher) erst bei der Ex Post Evaluierung erhoben werden, besteht – insbesondere bei einem Programm mit so langer Laufzeit wie A*plus*B - eine zu große Verzerrung aufgrund mangelnder Dokumentation oder fehlerhafter Erinnerungen.

Im Falle von AplusB wurde für die zeitgerechte Datenerhebung im obigen Sinne eine Kooperation mit den Initiatoren des Forschungsprojektes „Akademische Spin-off-Gründungen in Österreich" angedacht.[13]

Lessons learned

In nachfolgender bullet Point Liste soll nun die für die Evaluierung von A*plus*B bedeutenden Ergebnisse der konzeptionellen Vorüberlegungen, die Ausgangspunkt für das nachfolgende Kapitel sind, zusammengefasst werden.

- Um von einem rollierenden System der Evaluierung mit ausreichenden Feedback Schleifen sprechen zu können, ist eine Einpassung von Programmevaluierungsschritten in die bis jetzt festgeschriebenen Zentrenevaluierungen notwendig. Ehestmöglich sinnvoller Zeitpunkt für eine Programmevaluierung wäre 2006.
- Die erste Zwischenevaluierung nach drei Jahren soll stark auf ihre informationsorientierte Lernfunktion für die Zentren und das Programmmanagement zugeschnitten werden.
- Eine Legitimation des Einsatzes öffentlicher Mittel zur Förderung akademischer spin offs kann nicht von den Zentrenevaluierungen erwartet werden, sondern soll zentraler Auftrag an die Programmevaluierung werden.
- Gleiches gilt für typische Fragestellungen der Konzeptevaluierung wie jenen nach der Interventionslegitimation und einer Motivationsanalyse.
- Design- und Prozessevaluierung soll zentraler Bestandteil der Evaluierungsschritte auf Projekt- bzw. Zentrenebene sein.
- Die Aufgabenstellung an die EvaluatorInnen in der Prozessevaluierung ist so zu formulieren, dass es zu keinen Überschneidungen mit den Controllign-Tätigkeiten der TIG kommt.
- Im Sinne einer Evaluierungsökonomie kann es bei zwei aufeinander folgenden Interim - Evaluierungsschritten zu einer Abstufung der Fragestellungen zu den beiden Zeitpunkten kommen.
- Eine Wirkungsanalyse ist frühestens beim zweiten Evaluierungsschritt nach fünf Jahren möglich.
- Für eine umfassende Wirkungsanalyse ist für eine zeitgerechte Datenerhebung Vorsorge zu tragen.

[13] ZEW und Joanneum Research: „Akademische Spin-off-Gründungen in Österreich", Forschungsprojekt im Auftrag des bm:vit.

5 Die Evaluierungen in A*plus*B

Das Programmmanagement des A*plus*B Programms der TIG definiert als Ziel der Evaluierungen im gegenständlichen Programm, dass diese zeigen können sollen, dass öffentliche Investitionen in das Programm lohnen und eine entsprechende Transparenz der Mittelverwendung gewährleisten. Evaluationen sind sowohl auf Programm- als auch auf Zentrumsebene vorgesehen - Charakter und Funktion dieser jeweiligen Evaluationen sind unterschiedlich, die einzelnen Elemente sollen aber sinnvoll zusammenwirken. Die Grundsätze der Evaluation sind in den Programm-Richtlinien festgelegt (vgl. Abschnitt 9 der Richtlinien). Als Basis aller Evaluierungsschritte ist ein umfassendes Monitoring und Controlling der Zentren vorgesehen.[14]

5.1 MONITORING UND CONTROLLING DER ZENTREN

Monitoring und Controlling der Zentren hat mehrere Zielsetzungen: Einerseits eine Gebarenskontrolle, um die Verwendung öffentlicher Mittel entsprechend des im Fördervertrag vereinbarten Sinnes. Andererseits werden durch die Sammlung von Informationen im Monitoring und Controlling bereits für die Evaluierung relevante Daten generiert, die als Datenbasis sowohl für die Programm- als auch für die Projektevaluierung dienen kann. Generell sind zwei Arten der Informationsgewinnung zu unterscheiden. Einerseits bei den Zentren(managerInnen) selbst, andererseits bei den Gründungsvorhaben bzw. GründerInnen in den Zentren: Das den Zentren, den Gründungsvorhaben und den GründerInnen abverlangte Berichtswesen im Rahmen von A*plus*B besteht aus folgenden Elementen:

1. *Jahresberichte und Halbjahresberichte:* Die dem Programmmanagement der TIG zu liefernden Dokumente der Zentren sind Jahresberichte und Halbjahresberichte, wobei die Struktur der beiden Berichte gleich ist. Der Halbjahresbericht umfasst das erste halbe Jahr, der Jahresbericht das gesamte Jahr.
2. *Geschäftsbericht:* Die Erstellung eines jährlichen Geschäftsberichts ist nicht verpflichtend, wird allerdings empfohlen.
3. *Gründungsmonitoring*: Das Gründungsmonitoring, in den Vereinbarungen der Zentren mit den GründerInnen fixiert, besteht aus Datenerhebungen zum Zeitpunkt des Eintritts des Gründervorhabens in das Zentrum, einer Datenerhebung zum Zeitpunkt der Gründung und schließlich Erhebungen nach erfolgter Gründung auf vier Jahre.

Mit diesen Berichtswesen können verschiedene Informationsbedürfnisse gedeckt werden: Jenes der TIG als Programmmanager, jenes der das Programm kofinanzierenden Landesstellen (für beide hinsichtlich des Gebarens als auch des Erfolges), jenes des politischen Entscheidungsträgers, dem bmvit, und schließlich jenes der Zentren- und der ProgrammevaluatorInnen. Zentrale Funktion ist also der Nachweis der Mittelverwendung und damit die Basis für die Auszahlung der Förderungen. Darüber hinaus sammelt das Monitoringsystem aber auch Daten als Basis für laufende statistische Auswertungen und für die Evaluationen im dritten und im fünften Jahr.

Eine der Eigenschaften, die Evaluierungssysteme auszeichnen ist die Strukturierung und Zusammenführung von Informationsbedarf, den unterschiedliche Beteiligte aus einer unterschiedlichen Motivation heraus hegen. Die Vermeidung etwaiger Doppelgleisigkeiten - zum Beispiel die Erhebung von

[14] Fragebögen, Datenerfassungsbögen etc. sind auf der Homepage des AplusB Programms, www.aplusb.at, downloadbar.

Unternehmenskennzahlen zum jetzigen Zeitpunkt durch die TIG und zu einem späteren durch die EvaluatorInnen – ist wünschenswert, um Evaluierungs-Fatigue zu vermeiden und generell, um die Abwicklung eines Programms ressourcenschonend zu gestalten. Dieses Ziel vor Augen wurde im Falle von A*plus*B verschiedener Informationsbedarf zusammengeführt und in ein und die selbe Erhebungssystematik eingebracht.

Jahres- und Halbjahresberichte, Geschäftsberichte

Die Darstellung umfasst alle im Berichtszeitraum geplanten und / oder durchgeführten Aktivitäten. Die tabellarische Darstellung inklusive notwendiger Erläuterungen steht im Vordergrund, es wird an dieser Stelle jedoch auch die Möglichkeit zu einer „Selbstdarstellung" der Zentren gegeben werden. [15]

Gründungsmonitoring

Das Gründungsmonitoring sammelt Daten über die Gründungsvorhaben und GründerInnen - vor allem für die Evaluierung von Zentren und Programm, in zweiter Linie auch für die permanente Dokumentation der Bestände. Hierbei sind drei Arten der Datenerhebung bei den Gründungsvorhaben bzw. bei den GründerInnen zu unterscheiden: eine Datenerhebung beim Eintritt ins Zentrum als Gründungsvorhaben, eine Datenerhebung zum Zeitpunkt der Gründung, sowie eine Datenerhebung nach Austritt aus dem Zentrum (auf mindestens jährlicher Basis) bis 4 Jahre nach Austritt.

Rechnungswesen/Controlling

Die AplusB-Zentren sind als GmbHs eingerichtet und damit aufgrund der gesetzlichen Bestimmungen zur Führung eines Rechnungswesens nach den Bestimmungen des Handelsgesetzbuches (HGB) verpflichtet. Im Zuge der Halbjahres- und Jahresberichte sind von den Zentren auch Zahlen zum Rechnungswesen bzw. Erläuterungen abzugeben. Um den dafür notwendigen Aufwand zu minimieren, wurden die betreffenden Tabellen durch die TIG dem handelsrechtlichen Gliederungsschema angepasst. Die Erhebung erfolgt in zeitlicher Übereinstimmung mit dem Gründungsmonitoring.

5.2 DIE EVALUATION DES PROGRAMMS

In den Richtlinien des Programms ist eine Zwischen- und eine Endevaluation der Zentren durch externe GutachterInnen vorgesehen. Die Beauftragung der Programmevaluation sowie die Formulierung der Terms of Reference hat durch das verantwortliche Bundesministerium zu erfolgen, sie ist aber für die Schaffung eines Evaluierungssystems insofern von Belang, als dass durch Monitoring- und Berichtswesen auf Zentrenebene eine sinnvolle Informationsgrundlage für die Programmevaluation geschaffen werden kann.

[15] Im Zuge von Prüfungen vor Ort durch die TIG werden die genehmigten Zahlen, die Förderungsbestimmungen und die Gebarung in regelmäßigen Abständen kontrolliert. Die über die Gründungsvorhaben bzw. die GründerInnen eingeholten Informationen werden von Seiten der TIG stichprobenartig kontrolliert.

Tabelle 5: Überblick über die Evaluationen auf Zentrumsebene

	ex ante Bewertung	Erste interim Evaluation	Zweite interim Evaluation	End-Evaluation
Zeitpunkt („*Wann wird evaluiert?*“)	Vor Beginn	im 3. Jahr	im 5. Jahr	im 10. Jahr
Evaluations**gegenstand** („*Was wird evaluiert*?“)	Antrag	Zentrum	Zentrum Gründungen	Zentrum
Durchführende/r (*„Wer evaluiert?“ „Wer bewertet?“*)	BM auf Empfehlung einer ExpertInnenjury, Formalprüfung durch TIG	externe GutachterInnen	externe GutachterInnen	externe GutachterInnen
Konsequenz *(„Welche Folgen hat die Evaluierung?“)*	Go or Non-go	Empfehlungen	Stop or Go	noch fraglich

5.3 DIE EVALUATION DER ZENTREN

Ex-ante Auswahl der Zentren

Die Ex-ante Auswahl der Zentren ist in den Richtlinien des Programms genau geregelt: Die Anträge durchlaufen eine formale Prüfung von Seiten der TIG, die inhaltliche Prüfung erfolgt von Seiten einer eigens dazu eingerichteten ExpertInnenjury. Je nach Schwerpunktsetzung der Zentren ist zusätzlich eine Fachbegutachtung vorgesehen. Die ExpertInnenjury formuliert eine eindeutige Förderungsempfehlung an die/en zuständige/n MinisterIn, die wiederum entscheidet.

Evaluation im Dritten Jahr

Eine erste Zwischenevaluation ist im dritten Jahr vorgesehen. Diese Evaluation hat primär Empfehlungscharakter: es wird die Aufbau- und Managementarbeit sowie die Implementierung der vorgesehenen Maßnahmen des Zentrums bewertet, die Erfüllung der Auflagen durch die ExpertInnenjury überprüft, allfällige Probleme und Aufbauschwierigkeiten geortet. Thema ist auch der bisherige Zielerreichungsgrad laut Projektplan. Im Vordergrund dieses Evaluierungsschrittes steht der Lerncharakter.

Im dritten Jahr des Bestehens der Zentren können noch kaum Gründerkarrieren analysiert werden, daher ist der Focus der Evaluierung auf die Gründungsvorhaben anstatt die GründerInnen gerichtet. Wichtig ist eine qualitative, formative Bewertung[16] der gesamten Zentrumsaktivitäten. Die Evaluierung hat „Auditcharakter“ und wird unter internationaler Beteiligung durchgeführt. Basis für diesen Evaluierungsschritt nach drei Jahren sind die Auswahlkriterien, im Mittelpunkt steht die Evaluierung des gewählten Designs und der Prozesse. Von einer Wirkungsevaluierung wird zu diesem Zeitpunkt Abstand genommen.

[16] “The formative evaluator is more like the coach, not only passing judgement on performance but – especially helping the dancers understand how to improve their performance. […] understanding the dancing process is the central thing.” Erik Arnold, Ken Guy, Technology Difusion Programes and the Challenge for Evaluation. OECD 1997, Seite 70

Evaluation im Fünften Jahr

Im fünften Jahr nach Zentrumsstart findet eine zweite Zwischenevaluierung statt, die gleichzeitig ex-ante Evaluierung der Planungen des Zentrums für die verbleibenden fünf Jahre ist. Im Gegensatz zur Evaluation im dritten Jahr ist hier nicht nur die Arbeit des Zentrums sondern auch Zahl und Qualität der Gründungsvorhaben und Gründungen Gegenstand der Evaluation. Aufgrund dessen kommen unter anderem quantitative Indikatoren zum Einsatz, die im Rahmen des Monitoring- und Controllingsystems (vgl. hier das entsprechende Dokument der TIG zum "Berichtswesen") erhoben werden. Im Gegensatz zu der Evaluierung nach drei Jahren ist dies eine Evaluation mit sowohl qualitativen als auch quantitativen Elementen auf Basis des Monitoringsystems. Hier setzt dann auch eine Wirkungsanalyse an, Grundlage hierfür sind Auswahlkriterien und Programmziele.

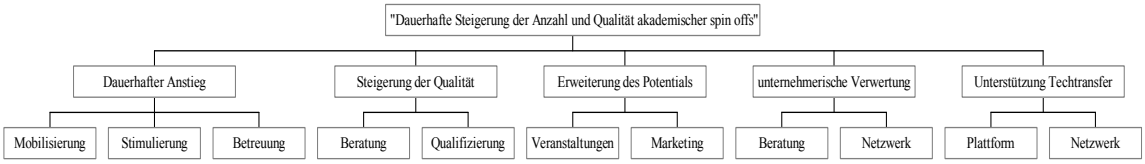
Abschließende Evaluation

Die Evaluation nach Ablauf des Förderzeitraumes überprüft schließlich die Tätigkeit des Zentrums aus einer ex-post Perspektive. Es wird hierbei wichtig sein, eine „Nach mir die Sinnflut"- Mentalität bei den Zentren zu vermeiden. Während bei der Evaluierung nach fünf Jahren eine Stop or Go Entscheidung getroffen wird, die eine gewisse Motivation des Zentrumsmanagement mit sich bringen wird, ist die Konsequenz der Evaluierung nach 10 Jahren noch unklar. Da bei einer Beendigung der Zentrumstätigkeit zu diesem Zeitpunkt die Motivation über die Vertragsverpflichtung hinaus kaum gegeben wäre, ist die Verknüpfung der Evaluation mit der Auszahlung der letzten Fördertranche zu erwägen. Darüber hinaus

5.4 ZIEL - INDIKATORENSYSTEM

Auf Seite 9 wurden detailliert die Ziele des Programms A*plus*B aufgeführt. Abbildung 6 gibt nochmals einen Überblick über Mission, Ziele und Aktivitäten in AplusB und knüpft so an die in Abschnitt 3.1 begonnene Logic Chart Analyse an. Zum Zielsystem in A*plus*B wurden verschiedene Indikatoren entwickelt[17], die klarmachen sollen, welche Maßstäbe an die Beurteilung der Zielerreichung gelegt werden.

Abbildung 6: Mission, Ziele, Aktivitäten



Diese Indikatoren wurden von der TIG veröffentlicht[18] um somit gegenüber den zu einem späteren Zeitpunkt zu evaluierenden Zentren vom Beginn an größtmögliche Transparenz zu garantieren. Informationsgrundlage für die Ermittlung der Indikatoren bzw. der relevanten Daten zu den Indikatoren ist das Berichtswesen (vgl. dazu Abschnitt 5.1).

[17] Im Zuge des Projektes „Grundsätze für ein Evaluierungssystem von A*plus*B" wurde vom Projektteam der Joanneum Research ein Katalog von möglichen Indikatoren entwickelt und an Monitoring- und Datenerfordernissen gespiegelt. In einem Workshop gemeinsam mit der TIG wurden aus diesem Katalog die weiter unten präsentierten Indikatoren ausgewählt. Kriterien für die Auswahl waren die Vergleichbarkeit mit Daten der spin off Erhebung.

[18] Konzept zur Evaluierung der AplusB Zentren, TIG Februar 2002, http://www.tig.or.at/bilder/154.pdf

Tabelle 6: Ziele & Indikatoren – eine Übersicht

Dauerhafter Anstieg	Qualität & Erfolg	Potential	Verbesserung der Verwertung	Techtransfer
- Anzahl der Gründungsvorhaben - Anzahl d. Gründungen	- Gründungen in High-Tech Branchen - Qualifikation der MitarbeiterInnen - Überlebensraten - Wachstumsraten - F&E Intensität - Forschungskooperationen	- Stimulierungsaktivitäten - Marketingaktivitäten - Haltung gegenüber Unternehmensgründung - Bekanntheitsgrad des Zentrums - "Entrepreneurship" in der beteiligten akademischen Einrichtungen	- Forschungsprojekte - Kooperation mit akademischer Herkunft - Gründungen, die auf der Verwertung akad. Forschungsergebnisse basieren	- Kooperation mit Transferstellen - Einbindung v. Unternehmen - Präsenz des Zentrums in der regionalen Wirtschaft - Nutzung anderer Förderungsmöglichkeiten

Den Anstoß zur Entwicklung von Ziel – Indikatorsystemen gab auch im Falle von AplusB die Notwendigkeit zur Transparenz – ProgrammteilnehmerInnen sollen von Beginn an darüber informiert werden, nach welchen Maßstäben ihre Arbeit beurteilt werden wird. Eine darauf aufsetzende , weiterführende Überlegung ist die der Verbindung von Ziel – Aktivität - Output – Outcome – Impact:

- Eine Zuordnung Ziel – Aktivität ist ein wesentliches Element einer Design- und Prozessevaluierung: Werden bestimmte Zielsetzungen vernachlässigt? Manche überverfolgt? Sind bestimmte Aktivitäten ressourcenmäßig unterausgestattet?
- Eine Unterscheidung Output / Outcome / Impact ermöglicht eine Zuordnung und Gewichtung von Indikatoren für verschiedene Evaluierungszeitpunkte: Outputmessung kann Gegenstand der Interimsevaluierungen (auch der nach drei Jahren) sein, Impactmessung nur in den seltensten Fällen.

Alle weitergehenden Überlegungen fallen in die Zuständigkeit des Evaluators bzw. der Evaluatorin, nicht in das Aufgabengebiet der Evaluierungssysteme.

Nachfolgend werden nun die im Gründungsmonitoring und im Berichtswesen der Zentren erhobenen Indikatoren zusammengefasst und nach Output- Outcome- und Impacteffekten gewichtet.

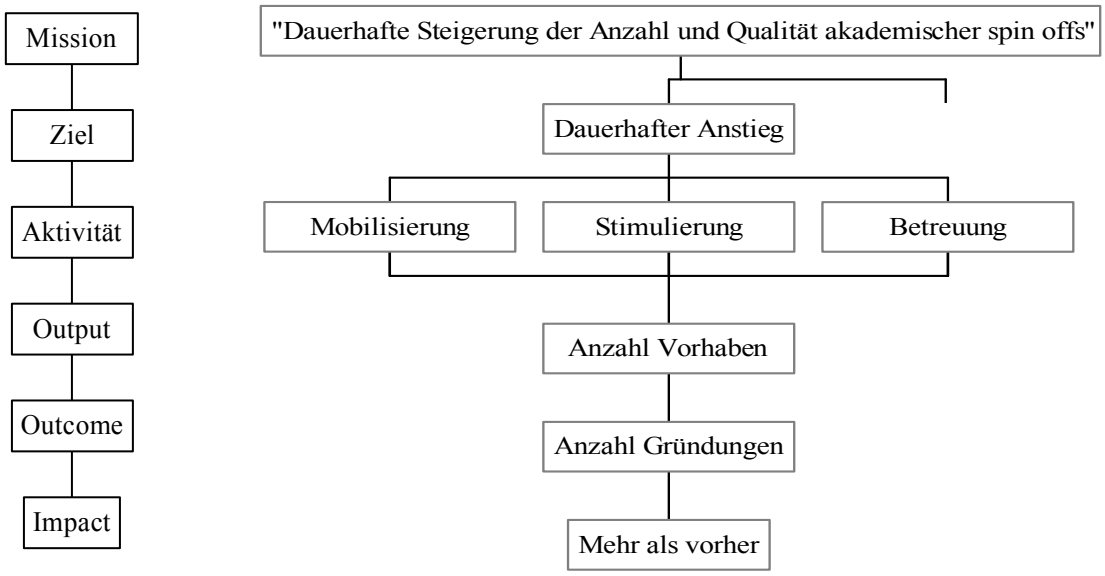
Dauerhafter Anstieg der Zahl akademischer Spin-offs

Indikatoren zum Ziel "dauerhafter Anstieg der Zahl akademischer Spin-offs" sind primär quantitativer Natur. Hierzu zählen:

- Anzahl der Gründungsvorhaben im Zentrum
- Anzahl der Gründungen aus dem Zentrum

Diese beiden Daten werden zu verschiedenen anderen Kenngrößen in Beziehung gesetzt (Spin-off Gründungen ohne Unterstützung durch ein A*plus*B-Zentrum, geschätztes Potenzial, generelles Gründungsaufkommen in der Region). Zeitpunkt der Überprüfung ist die Evaluierung im fünften Jahr.

Abbildung 7: Logic Chart „Dauerhafter Anstieg der Zahl akademischer Spin-offs“



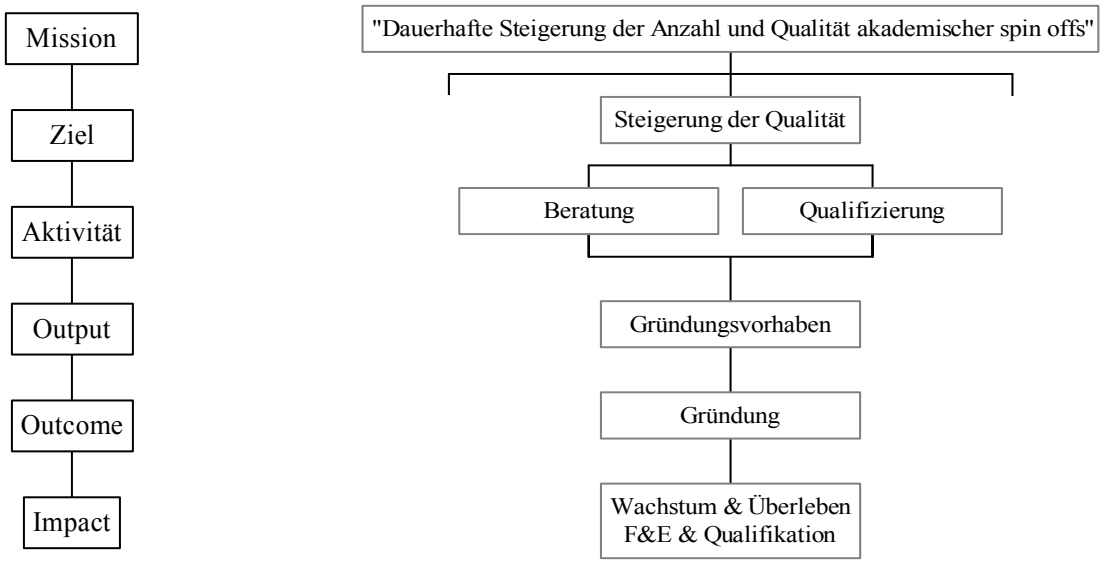
Steigerung der Qualität und Erfolgswahrscheinlichkeit akademischer Spin-offs

Indikatoren zum Ziel “Steigerung der Qualität und Erfolgswahrscheinlichkeit akademischer spin offs” sind wiederum primär quantitativer Natur. Hierzu zählen:

- Gründungen in High-Tech Branchen (ÖNACE Gliederung)
- Qualifikation der MitarbeiterInnen (höchster Bildungsabschluss)
- Überlebensraten (Bestands- bzw. Überlebenswahrscheinlichkeit)
- Wachstumsraten (Umsatz und Beschäftigte)
- F&E Intensität (F&E Ausgaben / Umsatz)
- Forschungskooperationen

Zeitpunkt der Überprüfung ist die Evaluierung im fünften Jahr.

Abbildung 8: Steigerung der Qualität und Erfolgswahrscheinlichkeit akademischer Spin-offs



Erweiterung des Potenzials an Unternehmensgründungen

Indikatoren zum Ziel „Erweiterung des Potentials an Unternehmensgründungen“ sind sowohl quantitativer als auch qualitativer Natur.

- Durchgeführte Stimulierungsaktivitäten (jeweils differenziert: welche Zielgruppen konnten erreicht werden)
- Marketingaktivitäten und genutzte Medien
- Bekanntheitsgrad des Zentrums bei AbsolventInnen und Studierenden (sowie gegebenenfalls anderer Zielgruppen)
- Positive Haltung gegenüber Unternehmensgründung bei den verschiedenen Zielgruppen
- Integration von "Entrepreneurship" in die Zielsetzungen und strategischen Planungen der beteiligten akademischen Einrichtungen

Zeitpunkt der Überprüfung ist die Evaluierung im dritten Jahr und im fünften Jahr.

Abbildung 9: Logic Chart Erweiterung des Potenzials an Unternehmensgründungen



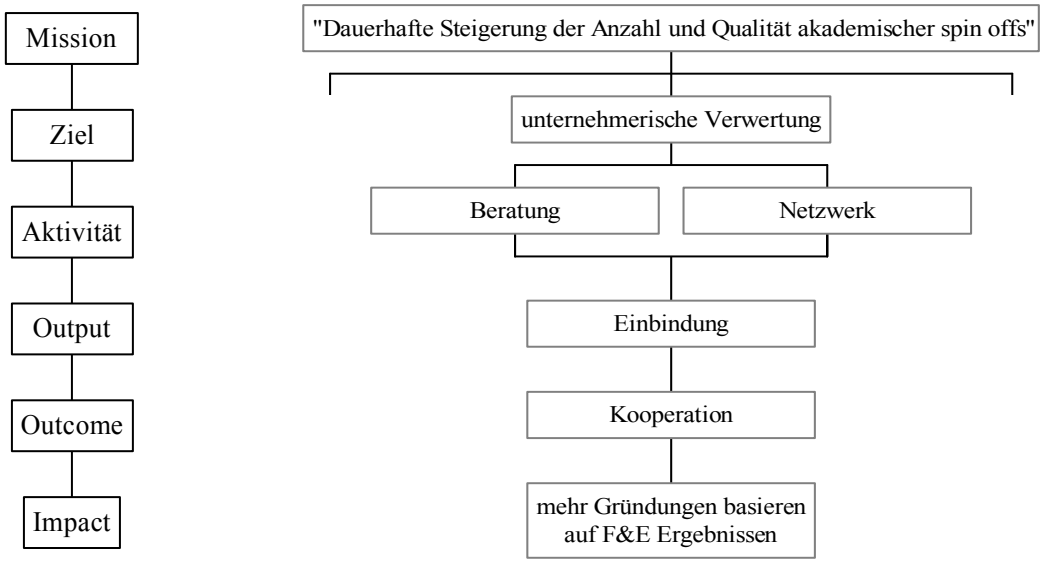
Verbesserung der Verwertung von Forschungsergebnissen

Indikatoren zum gegenständlichen Ziel sind primär quantitativer Natur, so wie:

- Einbindung der Gründungsvorhaben (im AplusB-Zentrum) in Drittmittelprojekte oder andere Forschungsprojekte an den akademischen Einrichtungen
- dauerhafte Kooperation der Gründer mit ihrer akademischen Herkunft (nach Verlassen des AplusB-Zentrums)
- Anzahl von Gründungen, die auf der Verwertung akademischer Forschungsergebnisse basieren (davon: Anzahl von Gründungen, die auf Patenten basieren)

Zeitpunkt der Überprüfung ist die Evaluation im fünften Jahr.

Abbildung 10: Verbesserung der Verwertung von Forschungsergebnissen



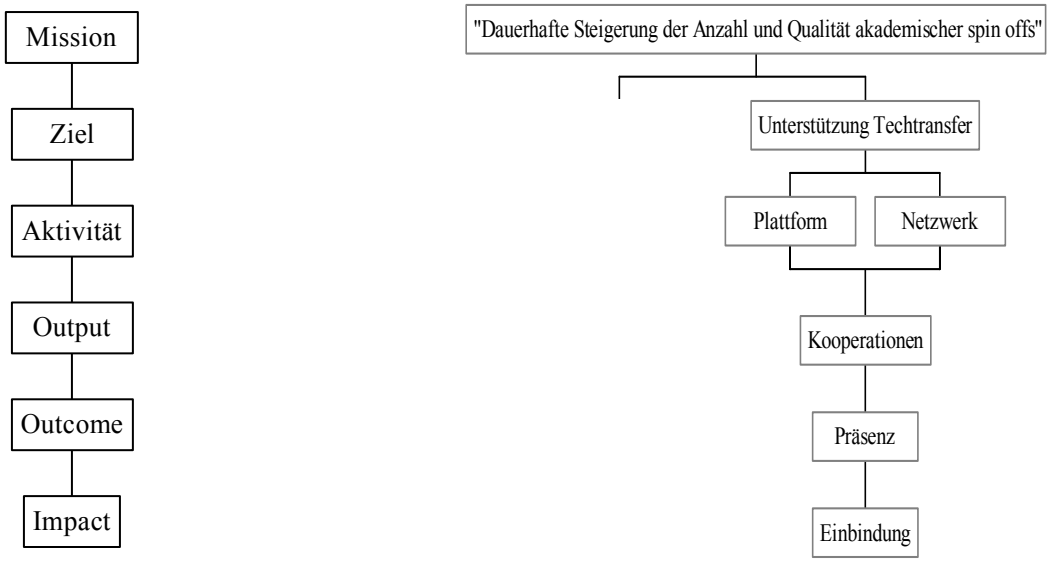
Unterstützung von Maßnahmen des Technologietransfers

Indikatoren zum gegenständlichen Ziel sind primär qualitativer Natur, so wie:

- Kooperation mit einer Transfer- bzw. Außenstelle der akademischen Einrichtungen
- Einbindung von Unternehmen in das Zentrum, Präsenz des Zentrums in der regionalen Wirtschaft
- Nutzung anderer Förderungsmöglichkeiten und Kooperation mit entsprechenden Förderungseinrichtungen

Zeitpunkt der Überprüfung ist die Evaluation im dritten Jahr und im fünften Jahr.

Abbildung 11: Unterstützung von Maßnahmen des Techtransfers



6 Anstatt eines Schlusswortes

Evaluierungssysteme sind nur dann sinnvoll,

- wenn sie den eigentlichen EvaluatorInnen genügend Freiheitsgrade in der Wahl der Methoden und in der Gewichtung der Fragestellung lassen.

Evaluierungssysteme sind kein Küchenrezept,

- den Programme, Programmrealtitäten, Methodenerfordernisse, Funktionen, Indikatoren usw. sind von Evaluierung zu Evaluierung verschieden.

Evaluierungssysteme sind ein interaktiver Prozess,

- die Güte eines solchen Systems ist von der Gesprächs- und Kooperationsbereitschaft von Policy Makern und von ProgrammmanagerInnen abhängig.

Evaluierungssysteme sind ein Mittel,

- Evaluierungs - Fatigue zu vermeiden. Wenn sich diese Müdigkeit allerdings darin erschöpft, dass Evaluierung als Instrument nicht ernstgenommen wird, scheitert jedes System.

Evaluierungssysteme berücksichtigen spätere Anforderungen einer Wirkungsanalyse, indem

- zu einem frühen Zeitpunkt Daten erhoben werden und somit die Güte und Qualität der Analyse gehoben wird.

Ein wesentliches Asset von Evaluierungssystemen ist die

- Strukturierung und Minimierung der Datenerhebungsschritte bei geförderten Unternehmen bzw. Personen.

Evaluierungssysteme sind ein Mittel, um

- Die idealtypische Einbettung von Evaluierung in den Politikzyklus sicherzustellen.

7 Bibliographie

BMVIT/ Technologie Impulse Gesellschaft: Sonderrichtlinien für die Förderung von Zentren zur Unterstützung akademischer Gründer und Gründerinnen AplusB Academia Business Spin-off Gründerprogramm 2001 bis 2006, Stand Juni 2001

ETAN Working Paper: Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes, Prepared by an independent ETAN Expert Working Group for the European Commission Directorate General XII, Programme Evaluation Unit 1999

Fritz O., Hutschenreiter G., Sturn D.: Evaluierung von FTE-Programmen: Zwischen Best Practice Entwicklung und österreichischen Strukturen. Joanneum Research, Wifo, Wien – Graz 1997

Fritz, O. et al.: Evina – Evaluierung von innovativen Aktionen in der Technologie-, Struktur- und Arbeitsmarktpolitik: Entwicklung von Methoden, Indikatoren und „Good Practice". InTeReg der Joanneum Research, Wien – Graz 2001

OECD Proceedings: Policy Evaluation in Innovation and technology. Towards Best Practices. OECD 1997

Plattform Forschungs- und Technologieevaluierung: Standards der Evaluierung in Forschungs- und Technologiepolitik. Wien 2003

Polt Wolfgang, Rojo, Jaime, Fahrenkrog Gustavo, Tübke Alexander, Zinöcker Klaus: "RTD Evaluation Toolbox", Socio Economic Evaluation of Public RTD Policies (EPUB) 2002

Schmid G., O'Reilly J., Schömann K.: Theory and Methodology of Labour Market Policy and Evaluation: An Introduction. In: Schmid G. et al.: International handbook of labour market policy and evaluation, Cheltenham [u.a.], Elgar, 1997

Sturn Dorothea, Zinöcker Klaus: How to make monitoring and evaluation match better? The case of the Austrian AplusB Programme. In: Plattform Newsletter 19: Evaluation & Monitoring of Programs and Projects, Wien 2003

Sturn, Dorothea (1999), 'Start auf dem Campus: Infrastruktur für Gründungen aus Hochschulen und Forschungseinrichtungen', unter Mitarbeit von Christoph Adametz, Markus Gruber, Franz Holzer, Mirjam Novakovic , Klaus Zinöcker, Projektbericht der Joanneum Research, Institut für Technologie und Regionalpolitik Wien – Graz.

TIG Technologie Impulse Gesellschaft: Academia Business Spin-off Gründerprogramm. Vorschriften für das Berichtswesen. Wien, Februar 2002

TIG Technologie Impulse Gesellschaft: Leitfaden zur Erstellung eines AplusB- Antrags, Wien im Juni 2001

JOANNEUM RESEARCH

INSTITUTE OF TECHNOLOGY AND REGIONAL POLICY

WORKING PAPER SERIES

Joanneum Research Working Paper Series

The Working Paper Series seeks to disseminate the results of research conducted within the Institute of Technology and Regional Policy of Joanneum Research to the broad academic community and other interested parties. Since much of the research is ongoing, the authors welcome comments from readers. Working Papers should not be reproduced without prior approval from the author. Electronic copies of the Working Paper Series can be found at: http://www.joanneum.at/rtg/wps For further questions, please contact: intereg@joanneum.at

JOANNEUM RESEARCH - *INSTITUTE OF TECHNOLOGY AND REGIONAL POLICY (INTEREG)*

VIENNA OFFICE:
WIEDNER HAUPTSTRAßE 76
A-1040 VIENNA, AUSTRIA
TEL: +43-1-581 75 20
E-MAIL: INTEREG@JOANNEUM.AT

GRAZ OFFICE:
ELISABETHSTRASSE 20
A-8010 GRAZ, AUSTRIA
TEL: +43-316-876 1488
E-MAIL: INTEREG@JOANNEUM.AT