

**Nützlichkeit und Nutzen der
Programmevaluationen im Bereich
der österreichischen FTI-Politik.
Metaevaluation der Programmevaluationen
2003-2014.**

Studie im Auftrag des Rats für Forschung
und Technologieentwicklung (RFTE)

Anhänge zum Endbericht
Dezember 2015

Günther Landsteiner

Anhang 1: Liste der analysierten Evaluationsberichte

Liste der 20 analysierten Programmevaluierungen			
Jahr	Titel	Auftraggeber	Evaluatoren (Institute)
2014	Impact Evaluation of the Erwin Schrödinger Fellowships with Return Phase	FWF	Fraunhofer ISI
2014	Zwischenevaluierung des Förderschwerpunkts Talente	BMVIT	KMU Forschung Austria
2013	Ex-post Evaluierung der Kompetenzzentrenprogramme K_plus und K_ind/K_net	BMVIT, BMWFJ	AIT, Joanneum Research, Manchester Institute for Innovation Research
2012	Zwischenevaluierung der Dienstleistungsinitiative des bmwfj	BMWFJ	Technopolis Group
2011	COIN Evaluierung	BMVIT, BMWFJ	Technopolis Group
2011	Elita – Evaluierung der FWF Programme Elise Richter und Hertha Firnberg. Endbericht	BMWF	Convelop
2011	Evaluierung des Programms FEMtech	BMVIT	convelop, Com.X Institut für Kommunikationsanalyse und Evaluation, GESIS Leibniz-Institut für Sozialwissenschaften/ CEWS
2011	Evaluierung des Programms uni:invent - Endbericht	BMWF, BMWFJ	Joanneum Research
2011	excellencia Evaluationsbericht	BMWF	IHS
2011	Programmevaluierung Headquarter-Strategie	BMVIT	Technopolis Group
2010	Evaluation des Pilotprogramms „Josef Ressel Zentren“	BMWFJ	Convelop
2009	Bericht zur Evaluation des Förderprogramms Sparkling Science	BMWF	AQA
2009	Programmevaluierung Bridge - Endbericht	BMVIT	Technopolis Group
2008	AplusB Academia Business Spinn-off Gründerprogramm - Zwischenevaluierung	BMVIT	inno Germany AG
2007	Zweite begleitende Evaluierung des „Impulsprogramm creativwirtschaft“ der aws	AWS	KMU Forschung Austria
2006	Interimevaluierung von protec 2002+	BMWA	KMU Forschung Austria, Fraunhofer ISI, ARC Systems Research
2006	Interimevaluierung der Österreichischen NANO-Initiative	BMVIT	Technopolis Group, NMTC
2005	Österreichisches Genomforschungsprogramm - Programmmanagerevaluierung	BM:BWK	TIA Consulting
2005	Zwischenevaluierung der RIF 2000 Regionale Impulsförderung	BMVIT	Technopolis Group
2004	Zwischenevaluierung Mikrotechnik Österreich	FFF	Joanneum Research

ANHANG 2: Factsheets zu den analysierten Evaluationsberichten

Evaluationsbericht Nr. 1	Beobachtungszeitraum: 2003 -2006	
Klassifikation	Einordnung	
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)	
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact	
Standard	Kommentar	Einstufung*
N1 Identifizierung der Beteiligten und Betroffenen	Es findet sich lediglich ein Hinweis, dass Interviews mit Programmverantwortlichen geführt wurden.	4
N2 Klärung der Evaluationszwecke	Es finden sich folgende Angaben: Aufschluss darüber geben, ob Adjustierungen des Programms vorgenommen werden sollten, und ob das Programm fortgesetzt, eingestellt, oder mit einer anderen Initiative zusammengeführt werden soll. Es wird auf instrumentellen Nutzen abgezielt. Genauere Fragestellungen, die den recht allgemein umrissenen Zweck weiter spezifizieren, sind nicht zu erkennen.	3
N4 Auswahl und Umfang der Informationen	Die Evaluation agiert, obwohl sie einen recht umfassenden Anspruch erhebt, mit einer eingeschränkten Datenbasis, indem sie sich weitgehend auf erhaltene Monitoringdaten stützt. Daneben werden ergänzende Daten mithilfe dreier Methoden erhoben, die zwar als gut geeignet, aber letztlich nicht ausreichend einzustufen sind. Die EvaluatorInnen weisen selbst auf eine weitere Methode hin, die gut geeignet gewesen wäre, das Evaluationsdesign anzureichern, deren Einsatz jedoch nicht beauftragt wurde. Sie weisen auch darauf hin, dass eine vollständige Einschätzung der Erreichung der verschiedenen Ziele des Programms anhand der verwendeten Daten nicht möglich ist. Ein guter Teil des Berichts gilt der Aufarbeitung des Fördergeschehens, so wie es anhand der verfügbaren Monitoringdaten nachvollzogen werden kann. Die Evaluation hat dabei streckenweise den Charakter einer Einschätzung der Evaluierbarkeit des untersuchten Programms. Der Bericht zeigt Züge einer Überprüfung, wie weit die Analyse mithilfe der vorhandenen Daten geführt werden kann. Hinsichtlich des Programmaufbaus wurde eine Logic Chart erstellt. Sie gibt klar strukturierten Aufschluss über Zuordnungen von Programmaktivitäten bzw. Mitteleinsätzen zu angestrebten Outputs und gesetzten Zielen. Es kann aus den herangezogenen Informationen bzw. eingesetzten Methoden jedoch nicht geschlossen werden, dass die Logic Chart zum Anhaltspunkt für die Auswahl benötigter Information zur Beleuchtung der gesamten angenommenen Wirklogik des Programms, oder bestimmter Abschnitte der Wirklogik, geworden wäre. Vorgänger-Untersuchungen anlässlich der Einrichtung des Programms werden erwähnt, werden aber nicht in Form einer vergleichenden Analysestrategie aufgegriffen. Das durch das Programm geförderte FTI-Geschehen auf Projektebene wurde nicht über die in Monitoringdaten enthaltenen, ausschnitthaften Angaben hinaus untersucht. In der Auseinandersetzung mit Outputs des untersuchten Programms werden Informationen über Aspekte der Programmumsetzung, die für die Erreichung von Output und für die Entfaltung von Wirkungen wesentlich sein konnten, nur ansatzweise und ausschnitthaft analysiert. Für eine gute Einschätzung der Wirkung auf die vom Programm angezielten FTI-Aktivitäten in Österreich insgesamt, und von Elementen der Wirklogik wie z.B. Awareness in der Zielgruppe, wären zusätzliche Daten notwendig gewesen, die nicht nur diejenigen Teile der Zielgruppe erfassen, die bereits im Programm	4

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	einen Antrag gestellt haben. Fakten zum Fördergeschehen stehen eher freie Diskussionen und Plausibilitätsargumente zur Programmgestaltung gegenüber. Für die Einschätzung der Relevanz wären aktuelle Daten über die konkrete Situation des angezielten FTI-Segments wünschenswert gewesen. Die Analyse des gemäß der angegebenen Evaluationszwecke für eine künftige Zusammenlegung in Frage kommenden Programms (vgl. G2) reicht nicht aus, um zu einer gut fundierten Einschätzung zu gelangen.	
N5 Transparenz von Werten	Die Evaluation stützt sich auf objektive Daten, die teilweise behandelt werden, als ob sie aus sich selbst sprechen würden, und teilweise mit Sichtweisen von Programmverantwortlichen und Motiven aus der Fachliteratur angereichert werden. Es wird Zielerreichung anhand programmeigener Zieldefinitionen untersucht. Im Umgang mit den Zielformulierungen kommen mehrere unterschiedliche Konzepte zur Anwendung, die – als Auslegungen des mit den Zielformulierungen Gemeinten und daher zu Überprüfenden - interpretativ an das Programm herangetragen werden. Nicht an allen Stellen kommt es hier zu einer eindeutigen Klärung der eingesetzten Konzepte. Stellenweise, aber keineswegs durchgehend, wird dieser unklare Status der Maßstäbe auch berichtet. Es wird auch nicht vermittelt, ob bzw. inwiefern eingesetzte Konzepte mit Programmverantwortlichen geklärt wurden, oder ohne solche Klärungen von den EvaluatorenInnen eingebracht wurden. In der Auseinandersetzung mit der Programmanlage werden an verschiedenen Stellen der Analyse Bewertungsmaßstäbe angegeben, die in Bezug auf bestimmte Fragestellungen eingesetzt werden. Es wird jedoch nicht in einer kompakten und leicht auffassbaren Weise berichtet, warum die Evaluation welche Maßstäbe an das Programm heranträgt, und wie sich diese zu einer Gesamteinschätzung fügen. Mögliche Alternativen zu eingesetzten Betrachtungsweisen werden nicht abgewogen. Es wird auch nicht erkennbar, dass klare Evaluationskriterien am Beginn der Untersuchung formuliert worden wären, um später erhaltene Ergebnisse einstufen zu können.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in einer gut verständlichen Sprache abgefasst und präsentiert wesentliche Informationen in übersichtlichen Diagrammen und Tabellen. Er ist jedoch in einer Weise gestaltet, bei der davon ausgegangen wird, dass BerichtsadressatInnen das untersuchte Programm, Kontextprogramme sowie im Berichtstext angezogene Quellen kennen. Die Verständlichkeit und Nachvollziehbarkeit für Nicht-InsiderInnen ist daher eingeschränkt. Eine Logic Chart versammelt übersichtlich wesentliche Programmelemente. Vergleicht man sie mit anderen, textlichen Angaben über das Programm, so ergibt sich freilich, dass doch keine klar strukturierte und vollständige Programmlogik vermittelt wird. Die Darstellungsweise ist insgesamt trotz eines auf den ersten Blick gut strukturierten Aufbaus nicht kompakt, sodass BerichtsadressatInnen Informationen, Argumente und Hinweise quer über Berichtsabschnitte hinweg zusammensuchen müssen und letztlich nicht sichergestellt ist, dass sie bei einer schnellen Lektüre alle wesentlichen Informationen zu einem Gesichtspunkt der Analyse jedenfalls erhalten bzw. zur Kenntnis nehmen. So muss etwa für eine vollständige Auffassung der Analyse von Zielerreichung zwischen verschiedenen Berichtsteilen hin- und her geblättert werden. Während wiederholt Ergebnisse anderer Untersuchungen referiert und Literaturangaben gemacht werden, bezieht sich der Bericht in manchen Argumentationen auf andere Evaluationsberichte in Insider-Manier, ohne den für die Argumentation ausschlaggebenden Inhalt wiederzugeben oder eine genaue Bezugsstelle anzugeben.	4
N8 Nutzung und Nutzen der Evaluation	Es finden sich keine Angaben, die darauf hindeuten würden, dass Schritte unternommen worden wären, um die Nutzung der Ergebnisse durch die AuftraggeberInnen vorzubereiten, oder neben	4

	ihnen andere Beteiligte oder Betroffene des Programms für die Evaluation und ihre Ergebnisse zu interessieren. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden sollen, werden nicht dargestellt. Deutliche Schwächen hinsichtlich mehrerer Standards beeinträchtigen die Nützlichkeit.	
F3 Vollständige und faire Überprüfung	Der Bericht benennt Stärken und Schwächen des untersuchten Programms. Es wird klar darauf hingewiesen, dass Schritte einer als sachdienlich erachteten Analyse wegen Datenmangels nicht durchgeführt werden konnten. Einschränkungen für die Analysierbarkeit von Zielerreichungen werden in allgemein-abstrakter Weise festgehalten. Eine detailliertere Darstellung, die alle BerichtsadressatInnen unmissverständlich auf diese Umstände hinweist und die Konsequenzen näher erläutert, wäre allerdings dienlich gewesen. Zur Einschätzung der Erreichung eines Programmziels besteht aufgrund interpretationsbedürftiger Formulierungen keine klare Herangehensweise (vgl. auch N5), und zugleich wird von den EvaluatorInnen auch Datenmangel für eine konsequente analytische Herangehensweisen argumentiert. Hier weisen die EvaluatorInnen auf mögliche alternative Sichtweisen hin, die sich bei einer anderen Vorgehensweise ergeben könnten. Letztlich treten allerdings Vermutungen an die Stelle einer datengestützten und nachvollziehbaren Analyse. An mehreren Stellen stützen sich die EvaluatorInnen bei der Benennung von Stärken auf Annahmen oder Vermutungen, ohne eine nachvollziehbare Analyse vorzulegen. Es wird bei mehreren Aussagen textlich nicht klar differenziert, inwiefern es sich um gesicherte faktengestützte Erkenntnisse handelt, und inwiefern um Annahmen oder Hintergrundwissen der EvaluatorInnen. Die Einschätzung der Zielerreichung gerät insofern zu einem Schwächen-Bericht, als minimale Abweichungen von Zielwerten mit deutlich stärkeren Abweichungen in der Interpretation gleichgehalten werden. Die Einschätzung der Relevanz des Programms gerät zum Stärken-Bericht, insofern in der Interpretation der herangezogenen Daten positive Hinweise die Oberhand gewinnen. Zu Umsetzungsaspekten enthalten sich die EvaluatorInnen einer Einschätzung, ob hier Stärken oder Schwächen vorliegen.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Die Evaluation stellt das untersuchte Programm, seinen Zielformulierungen gemäß, in den Rahmen des Portfolios derselben Agentur und untersucht das Verhältnis zu zwei der parallelen Programme. Das untersuchte Programm wird mit einem thematischen Programm einer anderen Agentur verglichen, für das eventuelle Überschneidungen speziell zu vermuten waren. Dies beschränkt sich jedoch auf eine Gegenüberstellung der Programmkonzeptionen. Die Evaluation steht unter dem Vorzeichen einer bevorstehenden Neugliederung der Agenturen und erbringt ihre zentralen Empfehlungen vor allem unter diesem Blickwinkel. Anknüpfungspunkte auf regionaler Ebene und Umsetzungsschritte gemeinsam mit Kooperationspartnern werden als Bestandteile der Programmlogik ausgewiesen. Regionale Aspekte des ausgelösten Fördergeschehens und regionale Initiativen werden in der Folge beschrieben, ohne allerdings auf vorhandene oder mögliche Synergien analytisch einzugehen. Kooperationsaktivitäten und deren Rolle für die Erreichung intendierter Outputs, Outcomes und Ziele werden nicht untersucht. Auf potentielle Synergien mit einer internationalen Fördermaßnahme wird hingewiesen, doch unterbleibt eine nähere	3

	Beschreibung und Analyse.	
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht enthält keine konzise und transparente Methodendarstellung. Es finden sich an verschiedenen Berichtstellen methodische Bemerkungen, doch reichen diese nicht aus, um die Analysestrategie gut nachzuvollziehen und die Tragfähigkeit aller herangezogenen Daten einschätzen zu lassen. Angaben zu den über die Monitoringdaten-Auswertung hinaus eingesetzten Methoden sind trotz Beigabe eines Fragebogens im Anhang unklar. Die Vorgehensweise der Evaluation ist trotz der diskursiven Verfasstheit des Berichts nicht leicht zu entziffern. Der Bericht gibt nicht in klarer Weise Evaluationszweck(e) oder Evaluationsfragestellungen an. Zusammenfassung und einzelne Berichtsabschnitte sind nach Fragen strukturiert, zu denen jedoch nicht angegeben wird, ob dies die ursprünglich gestellten Evaluationsfragen sind, die die Analyse angeleitet haben. Dies könnte einem Berichtscharakter geschuldet sein, bei dem offenbar vor allem an die AuftraggeberInnen gedacht wurde, die über die Evaluationsfragestellungen informiert sind. Der Bericht macht nicht ausreichend explizit, inwiefern das Programm anhand einer gegenwärtigen Sichtweise durch Programmverantwortliche untersucht wird, und inwiefern anhand der ursprünglichen Programmformulierung. Die vorhandenen Angaben reichen nicht aus, um die ursprüngliche Programmformulierung gut zu erkennen. Die Logic Chart wurde offenbar unter Bezug auf eine aktuelle Sichtweise des Programms zum Evaluationszeitpunkt erstellt, ohne dass dies klar kommuniziert würde. Während es einzelne Hinweise auf Einschränkungen der Analysierbarkeit gibt, wird bei der zentralen Diskussion von Zielerreichung keine systematische und gut erkennbare Auseinandersetzung damit angeboten, inwieweit Aussagen zum Evaluationszeitpunkt mithilfe vorhandener Daten getroffen werden können. Es kann nicht davon ausgegangen werden, dass die Evaluation in ihrer Vorgehensweise gut aufgefasst und deren Vor- wie etwaigen Nachteile gut erkannt werden können.	4
G8 Begründete Schlussfolgerungen	Schlussfolgerungen und Empfehlungen finden sich in Form eines von Ergebnispräsentationen getrennten Schlusskapitels, sowie jeweils am Ende von Kapiteln zu Datenanalysen. Diese über den Bericht verteilten Schlussfolgerungen und Empfehlungen sind nur bei einer Gesamtlektüre des Berichts auffindbar und mit den Inhalten des ausgewiesenen Schlussfolgerungskapitels kaum ident. Die EvaluatorenInnen weisen deutlich darauf hin, dass Schlüsse und Empfehlungen sich auf das untersuchte Programm beschränken und nicht auf andere Programme übertragbar sind. Hinsichtlich von Zielerreichungen kommt der Bericht weitestgehend zu transparenten, faktengestützten Schlussfolgerungen. Dass hier sehr knappe Abweichungen von Zielwerten genauso interpretiert werden wie deutlich stärkere Abweichungen, ist als Konzeptfrage zu deuten (vgl. N5). Das Ausmaß der Zielerreichung wird im Berichtstext und in den Schlussfolgerungen unterschiedlich bewertet. In anderen Hinsichten stützen sich die zentralen Schlussfolgerungen nicht ausschließlich auf eine transparente Datenlage, sondern auch auf Annahmen und Vermutungen der EvaluatorenInnen. Aus der identifizierten Tatsache, dass das Programm einen guten Teil der angedachter Umsetzungsschritte nicht gesetzt hat, wird keine Konsequenz gezogen. Der Bericht schliesst mit einem Aufzeigen von alternativen Handlungsoptionen für die Zukunft des Programms, wobei er Vor- und Nachteile benennt und auf jeweils schlagend werdende Konzeptions- und Umsetzungsaspekte hinweist. Für diesen Entwurf von Handlungsoptionen spielen konkrete Ergebnisse der durchgeführten Untersuchung nur eine untergeordnete Rolle. Wesentliche Teile der erbrachten Ergebnisse und im Text enthaltenen Schlussfolgerungen daraus werden hier nicht aufgegriffen.	4

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

Evaluationsbericht Nr. 2		Beobachtungszeitraum: 2003 - 2006	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Der Bericht macht keine Angaben dazu, ob und inwiefern Programmverantwortliche, weitere Programmbeteiligte, oder Betroffene in die Evaluation einbezogen wurden. Es ergibt sich aus einzelnen Berichtspassagen, dass Interviews mit nicht näher benannten Personen geführt wurden. Dies reicht für eine Einschätzung, welche Kreise auf diesem Weg von der Evaluation erfahren haben und sich in der Folge zumindest potenziell für ihre Ergebnisse interessieren konnten, nicht aus.	0	
N2 Klärung der Evaluationszwecke	Der Bericht macht keine klaren Angaben zu den verfolgten Evaluationszwecken. Er benennt einen Evaluationsgegenstand sowie teils Fragestellungen und teils Datenquellen als Grundlage der Vorgehensweise der Evaluation. Parallel dazu werden als Themen von Berichtsabschnitten (nicht als übergeordnete Zwecke) angegeben: Lernen für die Anpassung des laufenden Programms und für andere vergleichbare Programme. Schlüsse für die Zukunft des Programms ziehen. Diese Angaben entsprechen einer Ausrichtung auf instrumentellen und konzeptuellen Nutzen sowie Aufklärung. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären.	4	
N4 Auswahl und Umfang der Informationen	Die Evaluation stützt sich auf erhaltene Monitoringdaten und Programmunterlagen. Zusätzlich wurden offenbar Interviews im Bereich von Programmumsetzenden und Zielgruppen geführt, wozu allerdings genauere Angaben fehlen. An manchen Stellen werden nationale und internationale Berichte zur Beleuchtung von Einzelaspekten herangezogen. Obschon die Vorgehensweise als sachdienlich zu erachten ist, geht die Evaluation mit der vorgelegten Analyse jedoch deutlich über das hinaus, was durch die herangezogenen Datenquellen als mit gesicherten Belegen ausgestattet erachtet werden kann. Die Evaluation hat über weite Strecken eher den Charakter einer Peer Review bzw. eines Expertengutachtens, was ein grundsätzlich anderes Evaluationsmodell darstellt als eine Programmevaluation im Verständnis der Standards. Nach Maßstäben der Standards hätten zahlreiche weitere Daten herangezogen und Erhebungen angestellt werden müssen, um eine gesicherte Faktenbasis für Fragestellungen zu gewährleisten, die der Bericht behandelt. Dies betrifft Wirkungs- ebenso wie Umsetzungsaspekte (Prozessdaten). Eine systematische Analyse entlang der ausgewiesenen Programmziele erscheint anhand der verwendeten Informationen nicht möglich. Es werden zahlreiche Aussagen über verschiedenste Aspekte der Programmkonzeption, der Programmumsetzung, der Zielgruppe sowie des Programmkontextes getätigt, bei denen nicht klar ist, ob hinter ihnen Fakten stehen, zu denen die Belege nicht angegeben werden, oder ob persönliches Hintergrundwissen und Annahmen der EvaluatorsInnen zum Tragen kommen. Eine Logic Chart oder ähnliche Mittel zu einer strukturierten Gesamtauffassung der Programmelemente bzw. Programmlogik wurden nicht eingesetzt.	4	
N5 Transparenz von Werten	Die Evaluation stützt sich auf eine Programmbeschreibung und eine Aufarbeitung der Programmziele. Eine systematische Analyse in Bezug auf Zielerreichungen bildet jedoch nicht den Hauptfokus, eine klare Wahl als zentrale Bewertungsgrundlage kann folglich nicht angenommen werden.	5	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Erfahrungen und Sichtweisen in der Zielgruppe wurden, wo sie erhoben wurden, in einer intransparenten und teils wertenden Weise wiedergegeben und interpretiert. Auch dieser potenzielle Bewertungsmaßstab wird also nicht konsequent und adäquat angewendet. Bei der Identifikation von Schwächen und Verbesserungspotenzial stützen sich die EvaluatorInnen stark auf persönliches Hintergrundwissen und eigene Annahmen. Die EvaluatorInnen geben zugleich an verschiedenen Stellen in der Art der Formulierungen heraus klar zu erkennen, dass sie so vorgehen. Wie und warum bestimmte Aspekte der Programmanlage, des Programmgeschehens und des Kontexts, in dem sich das Programm bewegt, wie beleuchtet und kommentiert werden, wird nicht zum Gegenstand einer Erläuterung.	
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in allgemein-verständlicher Sprache abgefasst und frei von Technizismen. Tabellen und Grafiken, die anhand von Verzeichnissen auch leicht auffindbar sind, sowie grafische Textgestaltungen unterstützen die Informationsvermittlung. Ein Executive Summary ist nicht beigegeben. Der Bericht ist jedoch zugleich inhaltlich in einer Weise gestaltet, dass kaum davon ausgegangen werden kann, dass BerichtsadressatInnen, die mit dem Programm und dem Vorgehen der Evaluation nicht bereits vertraut sind, überall leicht folgen können. Es werden Darstellungen zum untersuchten Programm gegeben, die jedoch verstreut und jeweils ausschnittshaft sind. Eine kompakte Darstellung von Programmzielen liegt vor, doch finden sich in verschiedenen Berichtsabschnitten unterschiedliche Umgangsweisen mit unterschiedlich gewichteten Zielen. An manchen Stellen, teils auch mit schlussfolgerndem Charakter, drücken sich die EvaluatorInnen nicht so aus, dass jedenfalls zweifelsfrei von allen AdressatInnen verstanden wird, was sie meinen. Manche Daten werden in sehr detaillierter Weise dargestellt, ohne eine Interpretation anzubieten und so den BerichtsadressatInnen zu vermitteln, wie diese Daten aus Sicht der EvaluatorInnen zu verstehen sind. Schlussfolgerungen und Empfehlungen finden sich an zahlreichen Berichtsstellen, ohne als solche gut erkennbar ausgewiesen zu sein. Der unmittelbare Konnex zur vorangegangenen Argumentation erleichtert die Auffassbarkeit, doch setzt die Kenntnisnahme aller Schlussfolgerungen und Empfehlungen die genaue Lektüre des gesamten Berichts voraus.	4
N8 Nutzung und Nutzen der Evaluation	Das Ausmaß der aus der Berichterstattung entnehmbaren Einbeziehung von Beteiligten und Betroffenen (vgl. N1) lässt auf schlechte Voraussetzungen für eine Aufnahme der Evaluationsergebnisse unter Personen und Kreisen jenseits der unmittelbaren AuftraggeberInnen schließen. Die Annahme scheint zulässig, dass der Bericht für einen engen Kreis von HauptadressatInnen und SystemkennerInnen mit seinem Informationsangebot nützlich werden konnte. Zahlreiche Schwächen in Bezug auf verschiedene Nützlichkeits-, Fairness- und Genauigkeitsstandards schränken die Nützlichkeit ein. Es ist nicht auszuschließen, dass die AuftraggeberInnen einen Bericht in der vorliegenden Form auch erwartet haben. Da dies nicht gesichert ist, muss gemäß der Perspektiven der Standards eine Einstufung als wenig nützlich erfolgen.	4
F3 Vollständige und faire Überprüfung	In der Berichterstattung über Programmanlage, Programmkontext und erhobene Sichtweisen in der Zielgruppe wird an zahlreichen Stellen unmittelbar wertende Sprache eingesetzt. Bereits die einleitende Vorstellung der Programmkonzeption ist von einer deutlichen Skepsis gegenüber deren Sinnhaftigkeit gekennzeichnet. Für Erfahrungen und Bedarfslagen der Zielgruppe stehen an mehreren Stellen Annahmen der EvaluatorInnen ein. Berichtete Wirkungen auf Projektebene werden mit suggestiven	5

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Kommentaren versehen. Nicht alle Faktenlagen, die dargestellt werden, werden auch im konzeptuellen Zusammenhang interpretiert, und eine Begründung dafür wird nicht angegeben. Daten über die vom Programm Betroffenen (Antragsteller bzw. Fördernehmer) werden zum Teil in einer herablassenden Weise abgehandelt, die manchen BerichtsadressatInnen suggerieren könnte, dass diesen Betroffenen bzw. Begünstigten rationales Handeln im Sinne des untersuchten Programms nur bedingt zugetraut werden kann, bzw. dass die Glaubwürdigkeit der Zielgruppe des Programms in Frage gestellt wird.	
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Der Bericht ist um eine umfassende Kontextualisierung des untersuchten Programms bemüht. Das Programm wird in den Gesamtkontext von Innovationspolitik(en) auf Bundes- und Länderebene gestellt. Wesentliche Entwicklungen auf Bundesebene werden dargestellt und Daten auf Länderebene werden beigebracht. Es wird benannt, dass die Programmumsetzung auf Kooperationen auf Länderebene beruht, und dass kooperatives Verhalten der Fördernehmer eine zentrale Erwartung an die Wirklogik des Programms darstellte. Beidem wird jedoch nicht anhand aussagekräftiger Daten nachgegangen. Die aus dem Programm geförderten Projekte werden in den Kontext anderer, von den Förderempfängern durchgeführter Projekte gestellt. Die Analyse wird jedoch nicht auf eine Ebene geführt, wo der Beitrag und Stellenwert des untersuchten Programms für übergreifende Gesamtentwicklungen erkennbar würde. Ein Berichtsabschnitt ist einem Vergleich mit internationalen Programmen gewidmet. Während hier zahlreiche Eckdaten präsentiert werden, gelangt die Analyse nicht auf eine Ebene, auf der konkrete Wirklogiken bzw. Erfolgsfaktoren erkannt werden können.	3
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht gibt nicht in klarer Weise Evaluationszweck(e) oder Evaluationsfragestellungen an, denen bestimmte eingesetzte Methoden zur Beantwortung zugeordnet würden. Eingangs werden einige Fragestellungen genannt und auch deren Motivationen erläutert, doch hat dieses Set wenig Entsprechung in der weiteren Gestaltung verschiedener Berichtsabschnitte. Viel eher begegnen den BerichtsadressatInnen im Verlauf der Lektüre unterschiedliche Fragestellungen oder Datenbetrachtungen, die bei aller Reichhaltigkeit an Hinweisen nicht immer auf ein Gesamtkonzept beziehbar sind. Die Evaluation beschränkt sich nicht auf die Behandlung von Fragestellungen, die anhand von vorliegenden Daten gut eingeschätzt werden können, sondern bewegt sich weit darüber hinaus. Dabei wird nicht übersichtlich vermittelt, inwiefern Daten zur Beleuchtung welcher Fragestellungen vorliegen, und welche Einschränkungen für die Tragfähigkeit von Aussagen vorliegen könnten, die der Bericht tätigt. Der Bericht stützt einige Ergebnisse mit klaren Quellenangaben aus, referiert aber auch zahlreiche Daten, ohne dazu Belege anzugeben. Die berichtete Faktenlage wird somit über beträchtliche Berichtsstrecken nicht transparent oder hinsichtlich etwaiger Stärken oder Schwächen einschätzbar. Der Bericht enthält keine konzise oder vollständige Methodendarstellung und keine Dokumentation des Vorgehens der Evaluation. Er ist so abgefasst, als ob alle BerichtsadressatInnen bereits mit dem Vorgehen der Evaluation vertraut wären.	4
G8 Begründete Schlussfolgerungen	In einem getrennten Schlusskapitel werden Schlussfolgerungen und Empfehlungen präsentiert. Sie sind hinsichtlich ihres Bezugs auf berichtete Fakten kaum transparent und bringen mehrfach nochmals neue Annahmen bzw.	5

	Hintergrundwissen der EvaluatorInnen ein. Daten und Interpretationen, die in vorangegangenen Kapiteln berichtet wurden und für ein Lernen über das konkret erfolgte Programm wesentlich sein könnten, werden nicht aufgegriffen. Entsprechend dem Stil eines Expertengutachtens (vgl. N4) oder eines Konzeptpapiers geben die EvaluatorInnen teils spezifische Hinweise, teils bieten sie allgemein-abstrakte Überlegungen an. Im Sinne einer Programmevaluation, wie sie die DeGEval- und JC-Standards fassen, kann dies allerdings nicht als zureichend erachtet werden.	
--	--	--

Evaluationsbericht Nr. 3	Beobachtungszeitraum: 2003 - 2006	
Klassifikation	Kommentar	
Evaluationsrolle und -typ	formativ (Zwischenevaluation)	
Evaluationsschwerpunkte	Relevanz, Effektivität	
Standard	Kommentar	Einstufung*
N1 Identifizierung der Beteiligten und Betroffenen	Es wird über Workshops berichtet, die der Thematisierung der Programmanlage dienten. Da hierzu jedoch keine genaueren Angaben gemacht werden, kann nicht genau eingeschätzt werden, ob dabei die verschiedenen Programmverantwortlichen und mit der Umsetzung des Programms Betrauten vollständig einbezogen waren. Im Rahmen von Datenerhebungen wurden Programmverantwortliche, mit der Umsetzung des Programms Betraute und Programmteilnehmer einbezogen. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben konnten.	2
N2 Klärung der Evaluationszwecke	Angegeben wird: Überprüfen, ob gute Voraussetzungen für die erfolgreiche Umsetzung des Programms gegeben sind. Erfahrungen aus der bisherigen Umsetzung für die Weiterentwicklung des Programmdesigns und die Organisation der Umsetzung nutzbar machen. Gemäß diesen Angaben wird auf instrumentellen und konzeptuellen Nutzen abgezielt. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären.	3
N4 Auswahl und Umfang der Informationen	Evaluationsaufgaben, daran orientierte Evaluationsfragestellungen und methodische Herangehensweise werden klar und übersichtlich dargestellt. Die Art der Fragestellungen, der Methodeinsatz mit qualitativer Schwerpunktsetzung, und die zur Beantwortung der Evaluationsfragestellungen herangezogenen Informationen sind als durchwegs sachdienlich zu erachten. Die Programmstruktur wird unter Bezugnahme auf Programmdokumente entlang von Zielsetzungen und Umsetzungsschritten aufgeschlüsselt. Der Bericht setzt sich auf Basis der eingesetzten Methoden detailliert mit sinnvollen Fragen zur Programmumsetzung und zur Erreichung von Outputs und ersten Outcomes auseinander. Es wird ausdrücklich festgehalten, dass der Evaluationszeitpunkt für eine Abschätzung von Programmwirkungen zu früh ist, und die Evaluation beschränkt sich folgerichtig auf eine Analyse gesicherter Outputs und erster Outcomes. Eine Logic Chart oder andere strukturierte Aufarbeitung der Vorgehensweise des Programms in Bezug auf seine selbstgesetzten Ziele wird nicht erstellt. Wegen der komplexen Struktur des Programms, das mehrere Programmlinien, mehrere Auftraggeber, mehrere während der Programmlaufzeit restrukturierte Agenturen, und mehrere Steuerungsgremien involviert, kann dies als nachteilig erachtet werden. In der Folge ergeben sich an einzelnen Berichtsstellen auch gelegentlich Unklarheiten, welchen Status bestimmte beleuchtete Fakten in Bezug auf die Programmlogik haben und inwiefern sie Einfluss auf die Erreichung von selbstgesetzten Zielen haben können. Zu einer besseren Einschätzung der Erfassung und Mobilisierung der Zielgruppe(n) wäre, wie die EvaluatorInnen selbst festhalten, eine Datenbasis über die Gesamtaktivitäten in diesem FTI-Segment dienlich gewesen. Eine entsprechendes Datengewinnungsverfahren wurde nicht durchgeführt. Für eine gute Einschätzung der Sinnhaftigkeit von verschiedenen, teils erst geplanten Programmbestandteilen wäre es jedenfalls	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	günstig gewesen, nicht nur die an einer der Programmlinien bereits teilnehmenden Mitglieder der Zielgruppe(n) zu befragen. Hier wird implizit eine Identität von Bedürfnissen und Interessen zwischen Teilnehmern und Nicht-Teilnehmern unterstellt, die nicht zwingend Entsprechung in der Realität zu haben braucht. Eine echte Bedarfsanalyse kann hierdurch nicht ersetzt werden. Die Untersuchung der Relevanz hätte theoretisch auf eine bessere Grundlage gestellt werden können, wenn eine Erhebung zum aktuellen Marktpotenzial im Innovationssegment durchgeführt worden wäre. Während verschiedene verfügbare Unterlagen und Angaben der Zielgruppe zur Einschätzung der Relevanz des Programms herangezogen werden, fehlt so doch ein Bestandteil einer gut fundierten Gesamtperspektive. Ein Teil des Berichts widmet sich der Deskription des Fördergeschehens, ohne dass der Status dieser Daten zur Klärung von Evaluationsfragestellungen bestimmt würde oder die EvaluatorInnen Schlussfolgerungen daraus ziehen würden. Dies betrifft einerseits ein unter Umständen ungenutztes Potenzial dieser Daten für relevante Erkenntnisse. Andererseits hat sich dadurch die Evaluation nicht auf die Behandlung der wesentlichsten Gesichtspunkte zur Beantwortung ihrer Fragestellungen konzentriert.	
N5 Transparenz von Werten	Der Bericht wendet mit „Marktpotenzial“ ein wirtschaftliches Kriterium an, um die Relevanz des Programms einzuschätzen. Dies hat in der dargelegten Programmanlage seine Entsprechung, vereinseitigt allerdings doch die Betrachtung des sowohl auf Wissenschaft als auch Wirtschaft abzielenden Programms. Dies wird transparent gemacht, aber nicht näher begründet. An verschiedenen Stellen der vorgetragenen Analyse kommen daneben weitere andere Maßstäbe unkommentiert zur Anwendung. Der Bericht ist so verfasst und gegliedert, dass der Bezug zwischen Daten und Interpretationen und angewendeten Bewertungsparametern gut verständlich wird. Vorab definierte Evaluationskriterien, anhand derer späterhin Ergebnisse zu bewerten sein werden, werden nicht berichtet bzw. erkennbar eingesetzt. An mehreren Stellen, und in den zentralen Schlussfolgerungen und Empfehlungen, kommt allerdings ein bestimmtes Verständnis von Innovationsprozessen und darauf ausgerichteten Politiken zum Tragen, das wie selbstverständlich eingesetzt wird, ohne dass es aus der berichteten Programmformulierung heraus zwingend erschiene oder seine Wahl kommentiert würde.	3
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in einer klaren und strikt faktenorientierten Sprache abgefasst, die den Gehalt von Daten gut vermittelt und auch technische Sachverhalte allgemein verständlich transportiert. Wesentliche Ergebnisse werden in Tabellen und Grafiken dargestellt, die anhand von Verzeichnissen leicht aufgefunden werden können. Zur Vorstellung der Vorgehensweise der Evaluation und von Teilbereichen der Programmstruktur werden Grafiken eingesetzt. Ein Executive Summary ist beigegeben. Es erfolgt eine detaillierte Programmdarstellung, die allerdings in weiten Teilen keine Quellen ausweist und insofern eine gewisse Unsicherheit für die Nachvollziehbarkeit der vorgenommenen Untersuchung mit sich bringt. Bessere Quellenangaben hätten allen BerichtsadressatInnen klarer vermitteln können, inwiefern sie die Auffassungsweise des Programms als zutreffend erachten können. Der Bericht ist abgesehen von dieser Einschränkung gut geeignet, von verschiedenen Adressatenkreisen gut verstanden zu werden.	2
N8 Nutzung und Nutzen der Evaluation	Aus den Angaben über die Einbeziehung von Beteiligten und Betroffenen (vgl. N1) kann geschlossen werden, dass vor allem für Programmverantwortliche und –umsetzende gute Voraussetzungen geschaffen wurden, damit die Evaluation für sie nützlich werden konnte. Daneben wurden auch Programmteilnehmer auf die Evaluation aufmerksam gemacht, sodass sie sich für die Ergebnisse interessieren konnten. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden sollen, werden nicht dargestellt.	2

	Einschränkungen ergeben sich daraus, dass verschiedene Nützlichkeits-, Fairness- und Genauigkeitsstandards auch noch besser erfüllt werden hätten können.	
F3 Vollständige und faire Überprüfung	Der Bericht arbeitet sowohl Stärken als auch Schwächen des untersuchten Programms heraus. Dies wurde durch einen eigenen Schritt in der Vorgehensweise der Evaluation zur Kommunikation mit internationalen FachexpertInnen unterstützt. Auch in einzelnen Berichtsteilen wird die Interpretation von Ergebnissen nach identifizierten Stärken und Schwächen gegliedert. Ergebnisse werden so vermittelt, dass für die BerichtsadressatInnen erkennbar ist, welcher Grad an Aussagekraft und Tragfähigkeit ihnen von den EvaluatorsInnen zugemessen wird. Sowohl Sichtweisen der Programmverantwortlichen als auch Sichtweisen der Fördernehmer werden berichtet. Ein Konzept der wechselseitigen Spiegelung ist vorhanden, wird aber doch nicht in jeder Hinsicht durchgehalten. Die Programmdarstellung wird auch mit Bewertungen versehen, die offenbar Sichtweisen der einbezogenen Programmverantwortlichen darstellen, die an dieser Stelle direkt übernommen werden. In einem eigenen Schritt der Evaluation wurden internationale FachexpertInnen zur Identifikation von Stärken und Schwächen und Erarbeitung von Empfehlungen herangezogen. Die Ergebnisse dieses Schritts werden lediglich in einem Anhang wiedergegeben und es ist mit Ausnahme einer Textstelle nicht erkennbar, wie sie in den Diagnosen, Schlussfolgerungen und Empfehlungen der EvaluatorsInnen genutzt wurden.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Es wird auf den Kontext der Programmstehung eingegangen, sowie auf den internationalen Kontext für das angezielte Segment der FTI-Landschaft und die Motivation für dessen Förderung. Die Positionierung des Programms im österreichischen Förderportfolio wird thematisiert. Hier kommt es allerdings nur zu einer Einschätzung anhand weniger, teils auch selbst erhobener Daten. Dies ist sowohl im Zusammenhang mit Nichtverfügbarkeit von umfassenderen Datenbasen zu sehen, als auch mit einer Außerachtlassung von institutioneller Förderung (GUF). Die Analyse erreicht nicht eine Ebene, auf der erkennbar würde, welche Bestandteile des Förderportfolios in welcher Weise wirken. Zur Konstitution der Zielgruppe(n) und zu einigen Einflussfaktoren, die sich für die Wirksamkeit des Programms unter anderem aus dem Handeln anderer politischer Akteure ergeben können, werden Daten beigebracht. Diese Analyse hätte allerdings noch weiter vertieft und verfeinert werden können. Hier ist jedenfalls zu berücksichtigen, dass derartige Daten nicht verfügbar waren und nur mit recht beträchtlichem Aufwand erhoben werden hätten können.	3
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht enthält Angaben zu seiner Aufgabenstellung und zu Evaluationsfragestellungen sowie ein übersichtliches Methodenkapitel. Der Bericht enthält auch Angaben zur organisatorischen Vorgehensweise der Evaluation, die dazu beiträgt, ihre Einschätzbarkeit zu erhöhen. Eine Zuordnung der eingesetzten Methoden zu Fragestellungen oder Evaluationsschwerpunkten findet sich nicht. Es werden im Berichtsverlauf auch Fragestellungen behandelt, die in der Überblicksdarstellung nicht enthalten sind zu denen nicht mitgeteilt wird, ob es sich um ursprünglich vorgesehene Evaluationsfragestellungen handelt. Eingesetzte Methoden werden einerseits übersichtlich und so dargestellt, dass Vor- und Nachteile sowie die Tragfähigkeit der Ergebnisse gut eingeschätzt werden können. Es werden jedoch nicht alle eingesetzten Methoden in die Überblicksdarstellung einbezogen, und für manche der Methoden werden die benötigten	3

	Grundlagen für ihre Einschätzbarkeit nicht vermittelt. Die Methodendarstellung ist bei aller vorhandenen Qualität doch unvollständig, indem drei Schritte, auf deren Basis Evaluationsschwerpunkte behandelt und Schlussfolgerungen erbracht werden, hier keine Erwähnung finden. Die BerichtsadressatInnen werden im Berichtsverlauf teilweise mit Informationen bzw. Vorgehensweisen überrascht, die sie selbständig einordnen müssen. Darstellungen von Programmmerkmalen sind über verschiedene Berichtspassagen verteilt und nicht völlig konsistent. Die Evaluation wird anhand der Darlegungen nur bedingt als stringentes Gesamtkonzept erkennbar und hinsichtlich etwaiger Vor- und Nachteile ihrer Vorgehensweise einschätzbar.	
G8 Begründete Schlussfolgerungen	Schlussfolgernde Passagen finden sich am Ende einzelner Berichtsabschnitte, wo sie zwar im direkten Zusammenhang mit der vorangegangenen Datenpräsentation und Argumentation stehen, aber als solche nicht leicht erkennbar sind. Diese Schlussfolgerungen sind im Wesentlichen transparent in Bezug auf die dahinter stehenden Fakten bzw. in ihren Argumentationen gut nachvollziehbar. Es finden sich allerdings einige wenige Passagen, in denen der Bezug zum zuvor Berichteten nicht klar nachvollziehbar wird oder Argumente eingesetzt werden, die zwar nicht gänzlich unplausibel sind, aber doch auch in einer mehr abwägenden Weise dargelegt werden hätten können. Ein abschließendes Kapitel widmet sich, getrennt von den Ergebnisdarstellungen, nochmals zentralen Schlussfolgerungen und Empfehlungen. Hier wird die Programmanlage auch hinsichtlich der Positionierung des Programms verhandelt, wobei ein unmittelbarer Bezug zu vorhergehenden Darlegungen nicht immer gegeben ist. Es kommen teilweise Plausibilitätsargumente zur Anwendung, zu denen die EvaluatorenInnen nicht ausweisen, dass sie sich nicht direkt aus vorliegenden Daten ergeben, und dass andere Sichtweisen auch möglich sein könnten, z.B. wenn ein Zielkonflikt diagnostiziert würde. Der Berichtsabschnitt entfernt sich mit einigen Aussagen von den angegebenen Evaluationsschwerpunkten. Die abschließenden Schlussfolgerungen und Empfehlungen beziehen sich nur auf eine Programmlinie, doch wird dies nicht explizit gemacht. Während zum frühen Evaluationszeitpunkt nicht alle insgesamt vorgesehenen Programmlinien und –Maßnahmen untersucht werden konnten, werden einige, die durchaus Gegenstand der Analyse wurden, nicht aufgegriffen und in ihrer potenziellen Bedeutung für den Gesamterfolg des Programms vermittelt.	3

Evaluationsbericht Nr. 4	Beobachtungszeitraum: 2003 - 2006	
Klassifikation	Kommentar	
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)	
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact	
Standard	Kommentar	Einstufung*
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen von Datenerhebungsverfahren wurden mit der Programmsteuerung und -umsetzung betraute Personen sowie 9 Akteure der österreichischen FTI-Politik eingebunden. Des Weiteren wurden Fördernehmer des Programms auf diesem Weg eingebunden sowie Akteure eines Teilssegments der FTI-Landschaft, dem Bedeutung für die Realisierung von Programmserfolg zugemessen wurde. Eine Kurzbefragung richtete sich darüber hinaus an die Gesamtzielgruppe. Mit Programmverantwortlichen wurden zwei Workshops zur Diskussion von Zwischenergebnissen veranstaltet. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in die Planung der Evaluation einbezogen worden wären, um ihre Informationsbedürfnisse zu berücksichtigen.	2
N2 Klärung der Evaluationszwecke	Einleitend wird angegeben: Analyse des Programms und des bisherigen Programmverlaufs. Ausrichtung, Zielerreichung und Zielgruppenerreichung des Programms bewerten und etwaige Korrekturen im Hinblick auf die Weiterentwicklung empfehlen. In einem hinteren Berichtsabschnitt wird angegeben: Eine externe Bewertung der bisherigen Programmumsetzung vornehmen, die weitere Weichenstellungen für das Programm ermöglichen soll. Untersuchung der Positionierung im thematisch relevanten nationalen und internationalen Unterstützungssystem. Die vorfindlichen Angaben deuten somit auf eine Lernorientierung hin, sind dabei allerdings changierend und breit gehalten.	3
N4 Auswahl und Umfang der Informationen	Evaluationsfragestellungen, die die Vorgehensweise der Evaluation näher konkretisieren, werden nicht in umfassender Weise angegeben. Es werden vier Methoden parallel zum Einsatz gebracht, die angesichts der vorhandenen Angaben über die Themensetzungen der Evaluation grundsätzlich als gut geeignet erachtet werden können. Hinsichtlich der Programmwirkung werden erste Outcomes und Impacts auf der Ebene der geförderten Projekte analysiert, was jedenfalls eine sinnvolle Analysestrategie darstellt. Auf eine nur eingeschränkte Evaluierbarkeit angesichts des Evaluationszeitpunkts wird hingewiesen. Der Blickwinkel, der hier angelegt wird, kann als adäquat erachtet werden. Es werden jedoch nicht in systematischer Weise Daten herangezogen bzw. berichtet, die ein flächendeckendes Bild für alle Fördernehmer ergeben würden. Die Sinnhaftigkeit des Programmansatzes wird anhand internationaler Fachliteratur diskutiert. Für eine zentrale Begriffsklärung zur Programmanlage wird auf eine vorangegangene Evaluation im selben Themenbereich zurückgegriffen. In der Herangehensweise an das untersuchte Programm werden allerdings übergeordnete und operative Programmziele nicht kompakt und strukturiert, sondern an verschiedenen Textstellen dargestellt, und es wird im Berichtsverlauf nicht immer klar, wo sich der Bericht mit welchen Zielen bzw. welchen Mitteln, die das Programm zu deren Erreichung einsetzt, auseinandersetzt. Es findet sich eine Grafik zur Programmorganisation, eine Logic Chart oder ähnliche strukturierte Aufarbeitung der Gesamt-Programmlogik wurde jedoch nicht erstellt. Die vorliegende Analyse verknüpft so verschiedene detailorientierte und in Bezug auf das Programm doch aspekthafte Passagen. Die verschiedenen Programmlinien werden nicht gesamthaft aufgefasst. Umsetzungsaspekte können zum Teil nur schlecht hinsichtlich ihres Stellenwerts für den Programmserfolg	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	eingeordnet werden. Daten zur Bedarfsgerechtigkeit, Sinnhaftigkeit und Passgenauigkeit der Maßnahme wurden unter erfolgreichen Antragstellern (Fördernehmern) erhoben. Die Gesamtheit der Zielgruppe, auf die das Programm insgesamt abzielt, wurde dadurch nicht erfasst. Hinderungsgründe für die Erzielung von Programmserfolg können dadurch unter Umständen nicht erkannt worden sein. Die Evaluation behandelt an mehreren Stellen Gestaltungsaspekte des untersuchten Programms nur anhand der vorfindlichen Programmanlage im Allgemein-Prinzipiellen (z.B. anhand von Antragsvoraussetzungen), ohne an allen Stellen Daten heranzuziehen, durch die Passgenauigkeit oder Verbesserbarkeit dieser Programmanlage in einem realen Geschehen erkennbar werden könnte. Eine Erweiterung der durchgeführten Erhebungen in der Gesamtzielgruppe und in einem spezifischen Teilsegment um weitere Fragestellungen hätte die Vorgehensweise der Evaluation jedenfalls stärken können. Zur Erzielung einer noch aussagekräftigeren Analyse hätte es nicht unbedingt zusätzlicher Methoden bedurft, sondern eher einer anderen Ausgestaltung der eingesetzten Methoden im Verein mit einer strukturierten Aufarbeitung der Programmlogik. So können etwa mit Fragestellungen nach der Zufriedenheit von Fördernehmern mit der Zielrichtung und prinzipiellen Ausgestaltung von Programmlinien nicht Fragen abgedeckt werden, die Einflüsse auf Art und Ausmaß der Entfaltung des Programmgeschehens betreffen. Es wird unter anderem auch eine Fragestellung behandelt, die keinen erkennbaren Zusammenhang mit der Programmanlage hat. Grund dafür ist eine – von den EvaluatorenInnen auch benannte – rechtliche Vorgabe.	
N5 Transparenz von Werten	Der Bericht gibt als Bewertungsrahmen für das untersuchte Programm einerseits Zielerreichung und Zielgruppenerreichung an, und andererseits dessen Effektivität und Positionierung im thematisch relevanten nationalen und internationalen Unterstützungssystem. Der Ansatz der Zielerreichung und Zielgruppenerreichung wird allerdings nicht in systematischer Weise anhand durchgehend geeigneter Daten verfolgt. Zur Frage der Positionierung als Gesichtspunkt der Relevanz findet sich keine Angabe dazu oder Auseinandersetzung damit, wie hier dem Programm Wert zugewiesen werden kann bzw. soll. Eine Verankerung konkreter Betrachtungswinkel auf im Verlauf der Evaluation erzielte Beobachtungen wird nicht berichtet. Dies ist wohl auch im Zusammenhang damit zu lesen, dass ein Ausweis der verfolgten Evaluationsfragestellungen als Gesamtstruktur fehlt. Im Berichtsverlauf kommt es zum Einsatz unterschiedlicher „gängiger“ Bewertungsmaßstäbe wie Bekanntheit, Kundenzufriedenheit, Alleinstellungsmerkmal des Programms u.a.m., die jeweils wie selbstverständlich im Umgang mit einzelnen Ergebnissen argumentativ herangezogen werden. Sichtweisen, die im Rahmen von Interviews von Akteuren des FTI-Systems erhalten wurden, spielen darin eine wesentliche Rolle, was positiv zu werten ist. Ob an verschiedenen Stellen auch andere Betrachtungsweisen möglich gewesen wären, um dem Programm in seinen verschiedenen Gestaltungsaspekten Wert zuzuweisen, und inwiefern das Auswirkungen auf die Ergebnisse haben hätte können, wird nicht berichtet. Es wird davon auszugehen sein, dass vor allem BerichtsadressatInnen, die mit dem Programm und der Anlage der Evaluation nicht bereits vertraut sind, nur wenig Möglichkeit erhalten, sich im vielfältigen Umgang des Berichts mit verschiedenen Gestaltungsaspekten des Programms übergreifend zu orientieren.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in allgemein verständlicher Sprache abgefasst. Er setzt an manchen Stellen Tabellen und Grafiken ein, um Ergebnisse zu vermitteln, berichtet aber doch das Gros der Ergebnisse in rein textlicher Form. Eine stärkere Verwendung von Tabellen und Grafiken hätte die Auffassbarkeit an einigen Berichtsstellen weiter	4

	erhöhen können. Ein Executive Summary ist beigegeben. Am Anfang des Berichts steht eine kompakte Programmdarstellung, anhand derer auch mit dem Programm bislang nicht Vertraute einen Einstieg in die Berichterstattung der Evaluation finden können. Sie reicht allerdings für eine gute Verfolgbarkeit der Evaluation bzw. Auffassung ihrer Ergebnisse insofern doch nicht aus, als an verschiedenen Berichtsstellen immer wieder einzelne Gestaltungsaspekte und des Programms aufgegriffen werden, die in der Überblicksdarstellung nicht zu finden sind. An einzelnen Stellen berichten die EvaluatorenInnen erst nach einer Betrachtung von Daten, auf welchen Aspekt der Programmgestaltung sich diese beziehen, was jedenfalls die Auffassbarkeit deutlich beeinträchtigt. Im Zusammenhang mit einer fehlenden Aufschlüsselung von Evaluationsfragestellungen ergibt sich ein Text, der zwar in grobe Blöcke gegliedert ist, den die LeserInnen aber doch von der ersten bis zur letzten Seite genau durchgehen müssen, um zu einer Auffassung zu gelangen, wie die Evaluation vorgeht und was sie erbringt. Manche Ergebnisse werden zweimal berichtet.	
N8 Nutzung und Nutzen der Evaluation	Durch die Einbeziehung von Programmverantwortlichen und -umsetzenden sowie mehreren Akteuren des FTI-politischen Systems wurde ein guter Grundstein gelegt, damit die Evaluation für diese unterschiedlichen Akteure nützlich werden kann, bzw. Interesse an ihren Ergebnissen wecken konnte. Für die befragten Fördernehmer kann eine Weckung von Interesse ebenfalls noch angenommen werden, für die in Zusatzerhebungen mit punktuellen Fragestellungen adressierten Gruppen kaum. Der vorliegende Bericht kann mit den erbrachten Informationen und Darlegungen für die HauptadressatInnen (direkte AuftraggeberInnen) und InsiderInnen des FTI-politischen Systems nützlich sein, birgt hier allerdings doch auch Nachteile. Viele Informationen sind von einer Art, dass angenommen werden kann, dass gerade die Programmverantwortlichen sie bereits kennen. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden soll, werden nicht dargestellt. Es liegt ein insgesamt nicht gut nachvollziehbarer Bericht vor, der in seiner Gesamtanlage und seinem Vortrag doch einige Unklarheiten aufweist (vgl. N4, N5, N6, G3).	4
F3 Vollständige und faire Überprüfung	Der Bericht identifiziert anhand mehrerer unterschiedlicher Datenquellen, deren Ergebnisse jeweils zusammengezogen werden, in erster Linie Stärken des Programms. Einige Herausforderungen an das Programm werden ebenfalls benannt. Die zentralen Schlussfolgerungen sind nach identifizierten Stärken und Herausforderungen für eine Weiterentwicklung des Programms gegliedert. Qualitative Ergebnisse werden unter Beigabe ausführlicher Zitate dargestellt. Die Interpretation erfolgt allerdings öfters so, dass einzelne Zitate berichtet werden, die nicht durch weitere Hinweise begleitet sind, anhand derer eingeschätzt werden kann, welchen Stellenwert diese berichteten Aussagen im Gesamtbild aller in den Interviews erhaltenen Sichtweisen haben. Auch quantitative Ergebnisse werden in ganzen Berichtsabschnitten nicht umfassend und in Zahlen angegeben, was im Verhältnis zur angegebenen Evaluationsaufgabe der Auseinandersetzung mit Zielerreichung irritiert. Der Bericht übernimmt an zahlreichen Stellen die Sichtweise von Akteuren des FTI-politischen Systems. Dies wird auch explizit festgehalten und ist gemeinsam mit der Listung dieser Akteure in einem Anhang transparent. Da die EvaluatorenInnen sich an diesen Stellen allerdings nicht weiter äußern, bleibt unklar, ob sie diese Sichtweisen teilen und unterstützen, oder ob sie unter Umständen – auch unter Heranziehung anderer Daten - zu anderen Sichtweisen gelangt sein könnten. An verschiedenen Berichtsabschnitten werden Ergebnisse in schlussfolgernder oder wertender Weise präsentiert.	4

F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Es wird ein Überblick über die Entwicklung des Politikthemas auf Bundesebene unter Bezugnahme auf Fachliteratur gegeben. Sechs verwandte bzw. benachbarte Maßnahmen auf Bundesebene werden vergleichend dargestellt. Dieser Vergleich bezieht sich allerdings nur auf übergeordnete Programmziele und Eckdaten wie mobilisierte Fördersummen und gibt keinen Aufschluss darüber, auf welche Zielgruppen die Programm wie einzuwirken versuchen, und wo sie sich wie überschneiden oder ergänzen könnten. Es wird berichtet, dass das Programm regionale Komponenten und Synergiepotenziale hat, und eine Regionalverteilung des Fördergeschehens wird ausgewiesen. Weitere Daten, durch die regionale Voraussetzungen und Synergien beleuchtet werden könnten, wurden nicht herangezogen. Regionale Politikmaßnahmen werden nur oberflächlich skizziert. In internationaler Hinsicht wird das Programm neben mehrere ähnlich ausgerichtete Programme mehrerer Länder gestellt, und eine Teilnahme an einem internationalen Kooperationsprojekt wird erwähnt. Die Analyse konzentriert sich sodann auf die Benennung eines möglichen Gesichtspunkts für eine Anpassung des untersuchten Programms anhand eines herausgegriffenen internationalen Beispiels. Weitere Untersuchungen werden nicht angestellt. Es wird berichtet, dass das Programm durch eine Verteilung der Umsetzung verschiedener Programmlinien auf zwei Agenturen charakterisiert ist. Etwaigen Wirkungen dieser Kontexte für die Art und Weise, wie das Programm hier und unter Umständen doch in unterschiedlicher Weise seine Ziele zu erreichen versucht, werden nicht tiefergehend analysiert.	4
G3 Beschreibung von Zwecken und Vorgehen	Am Anfang des Berichts steht ein Methodenkapitel, in dem ein Überblick über die eingesetzten Methoden in ihrem Bezug zur Behandlung von Themenfeldern geboten wird. Angaben zu den zentralen eingesetzten Methoden erlauben eine grundlegende Einschätzung der Aussagekraft von berichteten Ergebnissen. Fragebögen oder Interviewleitfäden sind allerdings nicht beigegeben. Zwei eher untergeordnete Methoden, mithilfe derer einzelne Fragestellungen beleuchtet werden, werden nicht nachvollziehbar dargestellt. Die Vorgehensweise der Evaluation bei der Behandlung der Evaluationsschwerpunkte und -zwecke ist gleichwohl nicht leicht erkennbar. Sie wird nicht durch eine gesamthafte und übersichtliche Angabe von Evaluationsfragestellungen konkretisiert. Es wird keine konzise Korrespondenz zwischen Programmanlage und Einsatzpunkten der Evaluation hergestellt. Der Bericht wendet sich in seinem Verlauf unterschiedlichen Aspekten der Programmanlage und -umsetzung, erreichten Outputs und ersten Outcomes und Impacts sowie einigen Kontextfaktoren zu. Es bleibt letztlich jedoch schwer zu erkennen, welcher systematische Stellenwert verschiedenen Darlegungen zukommt. Es kann trotz vorhandener Angaben zu einigen relevanten Gesichtspunkten nicht davon ausgegangen werden, dass sich BerichtsadressatInnen - insbesondere solche, die mit dem Programm wenig vertraut sind und keine Expertise in FTI-Evaluation mitbringen - leicht zurechtfinden können, um die Vorgehensweise der Evaluation hinsichtlich von Vorteilen und etwaigen Nachteilen einschätzen zu können.	3
G8 Begründete Schlussfolgerungen	Ein abschließendes, klar getrenntes und leicht auffindbares Kapitel widmet sich Schlussfolgerungen und Empfehlungen. Die hier dargestellten Schlussfolgerungen unterhalten teilweise, aber nicht immer einen transparent gemachten Bezug zu dahinter stehenden Daten und Ergebnissen. Stellenweise kommt es zu Interpretationen, die nicht gut nachvollziehbar sind. Es werden teilweise Aussagen getroffen, für die in der Vorgehensweise der	4

	Evaluation mit den herangezogenen Informationen keine faktische Grundlage erarbeitet wurde. Auf der anderen Seite werden einige Hinweise auf Umsetzungsaspekte des Programms, die sich im zuvor Berichteten finden, nicht aufgegriffen. Ausgewiesene Schwerpunkte der Evaluation (vgl. N2) werden nicht bzw. nicht konzis und gut erkennbar beantwortet. Schlussfolgerungen finden sich auch in allen Einzelkapiteln, wo sie teilweise als Resümees ausgewiesen sind, teilweise aber auch nicht. An verschiedenen Berichtsstellen werden Ergebnisse in unmittelbar wertender Weise berichtet, ohne den Hintergrund dieser schlussfolgernden Präsentationsweise darzulegen.	
--	---	--

Evaluationsbericht Nr. 5		Beobachtungszeitraum: 2003 - 2006
Klassifikation	Einordnung	
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluierung)	
Evaluationsschwerpunkte	Relevanz, Effektivität	
Standard	Kommentar	Einstufung*
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen von Datenerhebungsverfahren wurden Programmverantwortliche, Betroffene des Programms (Fördernehmer und Nicht-Geförderte) sowie Stakeholder außerhalb des Programms einbezogen. In einem umfangreichen Prozess wurden zehn Akteursgruppen abgedeckt. Es wird berichtet, dass der Interaktion mit Programmverantwortlichen im Evaluationsprozess besondere Aufmerksamkeit geschenkt wurde. Auch die Tatsache, dass der Evaluationsauftrag in einem Verhandlungsverfahren definiert wurde, kann dazu beigetragen haben, die Informationsbedürfnisse des Auftraggebers gut geklärt werden konnten. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben konnten.	2
N2 Klärung der Evaluationszwecke	Angegeben wird: Programmdesign und Programmanagement untersuchen, um festzustellen, ob Verbesserungen in der Weiterführung des Programms notwendig sind. Klärung von zwei Hauptfragen: Soll das Programm fortgesetzt werden? Welche Verbesserungsmöglichkeiten sind zu identifizieren? Die Hauptfragen werden durch spezifische Evaluationsfragen differenziert. Es kann davon ausgegangen werden, dass die EvaluatorenInnen einen klaren Arbeitsauftrag hatten. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären.	2
N4 Auswahl und Umfang der Informationen	Die Hauptfragen der Evaluation werden in Form von spezifischen Evaluationsfragestellungen untersucht. Ein Mix aus 6 verschiedenen Methoden kommt zur Anwendung, der als gut geeignet eingeschätzt werden kann. Eine der Methoden kann als ungewöhnlich und innovativ bezeichnet werden. Eine Logic Chart wurde gemeinsam mit Programmverantwortlichen in einem Workshop entwickelt. Die in Form der Logic Chart aufgearbeitete Programmstruktur wird als der zentrale Schritt der Evaluation bezeichnet, der ihre gesamte Vorgehensweise informiert. Die EvaluatorenInnen geben an, dass mithilfe dieses Instruments Ziele und Grundannahmen des Programms sowie etwaige Inkonsistenzen und Lücken aufgearbeitet wurden. Dies ist hinsichtlich von Diskussionen von Programmaspekten im Berichtsverlauf nachvollziehbar. Dass die Aufarbeitung der Programmstruktur auch zum zentralen Anhaltspunkt für den Methodeneinsatz zur Beantwortung der Evaluationsfragestellungen geworden wäre, ist allerdings nicht nachvollziehbar. In der vorliegenden Form erscheint die Logic Chart nicht geeignet, Prozesse und Wirkungsweisen des Programmdesigns gezielt auf Beiträge zu Zielerreichungen hin zu untersuchen, oder die Rolle von potenziellen, in ursprünglich angenommene Wirkzusammenhänge intervenierenden Faktoren zu identifizieren. Die Auffassung der Programmstruktur hätte weiter gestärkt werden können, wenn auch eine Gliederung von übergeordneten und intermediären Zielsetzungen und differenziertere Zuordnungen der verschiedenen Instrumente und Maßnahmen dazu erfolgt wären. Auch eine Gliederung unterschiedlicher Outcomes und Impacts nach logisch-hierarchischen und zeitlichen Gesichtspunkten wäre sinnvoll gewesen.	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Zwei der eingesetzten Methoden bleiben letztlich unverbunden neben den anderen, tragenden Methoden stehen und ihre Ergebnisse werden nicht in Gesamtaussagen integriert. Es entsteht der Eindruck, dass die Evaluation entweder den konzeptuellen Stellenwert dieser Methoden nicht ausreichend bestimmt hat, oder nicht ausreichend Ressourcen zur Verfügung hatte, um die Integration zu vollziehen. Unter anderem wird versucht, den Stellenwert des Programms im Nationalen Innovationssystem zu beleuchten (Evaluationsfragestellung), doch werden hierfür benötigte bzw. sachdienliche Daten nicht herangezogen. Die vielfältige und differenzierte Untersuchung berührt verschiedentlich Punkte der Programmumsetzung, an denen es ihr weitere Daten erlaubt hätten, die Rolle identifizierter Faktoren noch genauer einzuschätzen, anstatt Plausibilitätsargumente heranzuziehen. Der Bericht enthält Hinweise, aus denen jedenfalls geschlossen werden kann, dass die EvaluatorenInnen einen hohen Kenntnis- und Bewusstseinsstand über evaluationsmethodische Anforderungen besitzen. Defizite der vorliegenden Untersuchung dürften daher in erster Linie Machbarkeits-Gründen im Rahmen des konkreten Auftrags geschuldet sein.	
N5 Transparenz von Werten	Eine explizite Darstellung zur Anwendung von Maßstäben, anhand derer dem Programm Wert zugemessen wird, findet sich nicht. Evaluationskriterien, mit denen vorab festgelegt wurde, wie späterhin Ergebnisse zu bewerten sein werden, werden nicht berichtet bzw. angewendet. An einigen, aber nicht allen Stellen sind einzelne Zielformulierungen des Programms leitend, ohne dass allerdings Zielerreichung im strikten Sinn überprüft würde. Sichtweisen von Betroffenen bzw. Stakeholdern spielen stellenweise eine tragende Rolle, allerdings ohne erkennbare übergreifende Systematik. Die Konsistenz der Programmanlage ist ein Bezugspunkt, es findet sich jedoch keine Auseinandersetzung damit, wie vollständig die Untersuchung diesbezüglich ist. Im Voranschreiten der Analyse kommt es unter anderem zur Anwendung von Maßstäben der Prozessqualität, der Zufriedenheit von Fördernehmern oder der Beobachtbarkeit von Mitnahmeeffekten, ohne dass dies vorgestellten Evaluationsfragestellungen direkt entspräche oder in anderer Form begründet würde. Die Analyse scheint streckenweise von einem Vertrauen getragen, dass die erhobenen und analysierten Daten aus sich selbst sprechen würden, und weist teils berichteten Daten keinen Stellenwert in der Identifikation etwaiger Vor- oder Nachteile des Programms zu.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist klar strukturiert und in einer flüssigen, nicht-technischen, stilistisch eher am Wissenschaftlichen orientierten Sprache gehalten. Zentrale Informationen werden in Tabellen und Grafiken dargestellt, die anhand entsprechender Verzeichnisse gut auffindbar sind. Wesentliche Aussagen werden im Text deutlich hervorgehoben. Ein Executive Summary ist in deutscher und englischer Sprache beigegeben. Das Programm wird mittels einer Logic Chart greifbar. Sie gibt allerdings nur einen beschränkten bzw. verkürzten Eindruck, der durch verschiedene Textpassagen notwendiger Weise ergänzt und erweitert wird. Während der Bericht sich auf das Programm in detaillierter Weise einlässt, ist durch die Verstreutheit von Angaben dem Gesamtzusammenhang des Programms und den Einsatzpunkten von dessen Untersuchung nicht immer leicht zu folgen. Nachteile der Berichterstattung über eine komplexe Untersuchung eines komplexen Programms können darin erblickt werden, dass der Bericht weder nach Evaluationsfragestellungen noch nach Programmelementen oder –zielsetzungen gegliedert ist. Es kann davon ausgegangen werden, dass nicht alle BerichtsadressatInnen leicht erkennen können, was wo wie genau untersucht wird (vgl. G3).	2
N8 Nutzung und Nutzen der Evaluation	Aus der nachvollziehbaren Einbeziehung mehrerer Akteursgruppen (vgl. N1) kann geschlossen werden, dass die Evaluation gute	3

	Voraussetzungen geschaffen hat, um für mit dem Programm Befasste ebenso wie für verschiedene Betroffene und Stakeholder nützlich zu werden und deren Interesse an den Ergebnissen zu wecken bzw. zu fördern. Aus der beigegebenen Liste der InterviewpartnerInnen wird allerdings ersichtlich, dass höhere Verantwortungsebenen im FTI-politischen System nicht einbezogen wurden. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden soll, werden nicht dargestellt. Einige Einschränkungen für die Nützlichkeit ergeben sich aus einer letztlich doch unvollständigen Untersuchung (N4, G2), wobei allerdings die Komplexität des Evaluationsgegenstands berücksichtigt werden muss, sowie einer nicht optimal nachvollziehbaren Analysestrategie (G3) und einer intransparenten Umgangsweise mit der Bewertung (N5).	
F3 Vollständige und faire Überprüfung	Die EvaluatorInnen berichten neutral über ihre Beobachtungen und teilen in klar getrennten Passagen mit, was sie zu diesen Beobachtungen denken und zu welchen Schlussfolgerungen sie kommen. Im Sinne der angegebenen Evaluationszwecke verfährt der Bericht kritisch-überprüfend, ohne dass er deshalb als einseitiger „Schwächen-Bericht“ bezeichnet werden könnte. Die EvaluatorInnen bringen an verschiedenen Stellen klar zum Ausdruck, zu welchen Meinungen sie hinsichtlich verschiedener Aspekte des untersuchten Programms gelangt sind. Manche Ergebnisse bleiben allerdings als solche stehen und die EvaluatorInnen geben nicht zu erkennen, ob sie hier Stärken oder Schwächen des Programms sehen. Zahlreichen Stakeholdern aus unterschiedlichen Gruppen wurde die Gelegenheit gegeben, sich zum Programm zu äußern. Sichtweisen verschiedener Gruppen auf das Programm werden allerdings jeweils isoliert in Bezug auf einzelne Programm- bzw. Analyseaspekte berichtet, sodass ein Gesamtbild der Positionierungen nicht zustande kommt. Die EvaluatorInnen nehmen diese Sichtweisen in unterschiedlichem Maß in ihre Argumentation auf bzw. verleihen ihnen zu unterschiedlichen Untersuchungsaspekten unterschiedliches Gewicht, wobei nicht immer transparente Begründungen mitgeliefert werden.	2
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das Programm wird in mehrfacher Hinsicht kontextualisiert: Die Rollen von Partnerorganisationen, die in die Programmumsetzung involviert sind, wird dargestellt und untersucht, wobei sich die Analyse allerdings auf Einzelaspekte beschränkt und als unvollständig bezeichnet werden kann. Das Programm wird in den übergreifenden Rahmen der Entwicklung von Schwerpunktsetzungen der nationalen FTI-Politik eingeordnet. Eine Untersuchung der Positionierung, aus der sich Schlüsse über die Art des Beitrags ableiten ließen, den das untersuchte Programm im Verhältnis zu anderen Programmen für das Innovationssystem leistet, wurde nicht angestellt. Internationale Vergleichsprogramme aus drei Ländern werden herangezogen, um – wie die EvaluatorInnen es ausdrücken - als Inspirationsquelle für Gestaltungsoptionen zu dienen. Für eine Programmlinie werden neun internationale Vergleichsprogramme skizziert. Die Zusammenstellung von Eckdaten reicht jedoch nicht aus, um die genauere Wirkungsweise der Programme in ihren Kontexten zu verstehen. Hier wären jedenfalls umfangreiche Recherchen notwendig gewesen, und es kann davon ausgegangen werden, dass Informationen auf dem eigentlich benötigten Niveau über weite Strecken gar nicht existiert haben dürften. Der Bericht weist an einzelnen Stellen auf Mechanismen im Umfeld des Programms bzw. in der Zielgruppe hin, die für die Erreichung von übergeordneten Programmzielen wichtig sein können. Eine systematische Analyse auf diesem Niveau wird	4

	jedoch nicht durchgeführt.	
G3 Beschreibung von Zwecken und Vorgehen	Die Hauptfragen der Evaluation werden anhand von spezifischen Evaluationsfragestellungen verfolgt, die einleitend dargestellt sind. Die Erfassung des Evaluationsgegenstands in sachlicher und zeitlicher Hinsicht wird klar dargestellt. Ein eigenes Methodenkapitel stellt die eingesetzten Methoden und das Rationale ihres Einsatzes übersichtlich dar. Im Bericht und in Anhängen finden sich ausreichende Angaben, um die Tragfähigkeit der erbrachten Ergebnisse einschätzen können. Nur mit Einschränkungen gilt dies allerdings für die innovative Methode. Die Logic Chart, die laut Bericht in Erarbeitung und Einsatzweise dem State-of-the-Art entspricht, bleibt allerdings sowohl gegenüber der textlichen Programmdarstellung als auch der Methodendarstellung isoliert. Sie verzeichnet auch Programmergebnisse, was einem Charakter als Leitmedium der Analyse des zu untersuchenden Programms wenig entspricht, und lässt sich schwer mit den – ebenfalls getrennt angegebenen – Evaluationsfragestellungen in Verbindung bringen. Es finden sich Berichtsabschnitte, in denen Programmaktivitäten behandelt werden, die in der Logic Chart nicht verzeichnet sind. Für BerichtsadressatInnen ist nur schwer nachzuvollziehen, welche Aspekte des Programms (Zielsetzungen, Umsetzungsschritte einschließlich von Maßnahmen jenseits des Förderungsvergabe) wie intensiv analysiert werden. Detaillierten Analysen an bestimmten Punkten stehen oberflächlichere Betrachtungen zu anderen Aspekten gegenüber, ohne dass dies als solches ausgewiesen oder mit einer Begründung versehen wird. Es wird nicht ausreichend transparent gemacht, welche Ergebnisse aus welchen Methoden welche Evaluationsfragestellungen beantworten lassen sollen. Es ist für BerichtsadressatInnen in den verschiedenen Kapiteln nicht leicht erkennbar, wo, wie und ob Evaluationsfragestellungen beantwortet werden. Eine klar strukturierte Beantwortung der Fragestellungen findet im Schlussfolgerungs-Kapitel statt. Um diese Beantwortungen nachzuvollziehen und einzuschätzen, müssen die BerichtsadressatInnen zuvor den gesamten Bericht studiert haben.	3
G8 Begründete Schlussfolgerungen	Ein von Ergebnisdarstellungen klar getrenntes Kapitel gilt Schlussfolgerungen und Empfehlungen. Die Schlussfolgerungen sind klar nach den Evaluationsfragestellungen strukturiert. Sie sind differenziert und abwägend formuliert und weitestgehend aus der vorangegangenen Analyse gut begründet. Zusätzlich geben die EvaluatorsInnen eine Reihe von Empfehlungen ab, die sich nachvollziehbar auf verschiedene Eindrücke stützen, die sie im Verlauf der Untersuchung gewonnen haben. Diese Empfehlungen sind großteils spezifisch und umsetzbar, und teils allgemein-prinzipieller Natur. Sie stellen allerdings keinen expliziten Bezug zu dahinter stehenden beobachteten Fakten her. An mehreren Stellen werden neue Fakten eingebracht. Es wird nicht zum Ausdruck gebracht, ob und inwiefern die verschiedenen Fragestellungen mit unterschiedlicher Sicherheit beantwortet werden können. Auch am Ende der Ergebniskapitel stehen hervorgehobene, leicht erkennbare Schlussfolgerungen und Empfehlungen zu den jeweils vorangegangenen Analysen, die in ihrem Bezug auf zuvor Berichtetes transparent sind. Ein Subkapitel ist mit „Schlussfolgerungen“ betitelt, berichtet allerdings neue relevante Fakten, um sodann getrennt transparente Schlussfolgerungen daraus zu ziehen. Der etwas unglücklich gewählte Titel mag im ersten Moment verwirren, kann aber nicht als substanzielle Beeinträchtigung der Erfüllung des Standards betrachtet werden. Eher bedenkenswert erscheint, dass hier nicht der Gehalt aller vorangegangenen Abschnitte hinsichtlich des Wesentlichsten zusammengeführt wird. Einzelne Schlussfolgerungen im Berichtsverlauf haben keinen transparenten Bezug zu Evaluationsfragestellungen oder zu berichteten Daten und verlassen sich auf Plausibilitätsargumente.	2

Evaluationsbericht Nr. 6		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	summativ (Ex post-Evaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Programmeigentümer, Programmgestaltende und ein Teil der vom Programm Betroffenen (Antragsteller) wurden im Rahmen von Datenerhebungsverfahren einbezogen. Anhand der Angaben des Berichts kann nicht eindeutig nachvollzogen werden, wie umfassend bzw. vollständig Programmverantwortliche und Umsetzende des Programms, das mehrere Programmeigentümer vereint und in der Gestaltung und Umsetzung Berührungspunkte mit anderen Programmen der betrauten Agentur aufweist, einbezogen waren. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben konnten. Eine Einschränkung ergibt sich aus einer Beobachtung, die unter G8 dargestellt wird.	3	
N2 Klärung der Evaluationszwecke	Angegeben wird: Den Programmverlauf reflektieren und Schlussfolgerungen und Empfehlungen für die Weiterentwicklung des Programms erarbeiten. Die Konzeption des Programms, seine Umsetzung, die Zielerreichung und feststellbare Wirkungen analysieren. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären. Im Fall des vorliegenden Programms wird hier in erster Linie an FTI-politische Verantwortungsebenen in mehreren Ressorts zu denken sein, während potenzielle Begünstigte eher nicht in Frage kommen.	3	
N4 Auswahl und Umfang der Informationen	Am Anfang des Berichts werden die Evaluationsaufgaben beschrieben und genauere Evaluationsfragestellungen aufgelistet. Die angestellte Untersuchung stützt sich auf mehrere Methoden, die grundsätzlich gut geeignet erscheinen. Die Ergebnisse mehrerer Methoden werden zusammengezogen, um einzelne Sachverhalte aus mehreren Blickwinkeln zu beleuchten (Triangulation). Dies stellt allerdings keine durchgehende Strategie für alle Analysepunkte dar. In mehrerer Hinsicht wäre allerdings eine Heranziehung zusätzlicher Informationen sachdienlich gewesen: • Programmwirkungen auf der Ebene von Kooperationspartnern, die auch in einer Evaluationsfrage als (indirekte) Zielgruppe des Programms verstanden werden, werden nicht untersucht. Dadurch ergibt sich eine eingeschränkte Analyse, in deren Rahmen die Gesamtlogik des Programms nicht vollständig eingeschätzt werden kann. • Ebenso hätte die substanzielle Behandlung der genannten Evaluationsfrage zusätzlicher Daten bedurft. • Teile der Zielgruppen, die sich am Programm nicht beteiligt haben, werden nicht untersucht. Zwei Evaluationsfragen bedürften allerdings zu einer guten Beantwortbarkeit einer umfassenden Analyse der Gesamtzielgruppe. Die eingeschränkte Vorgehensweise erscheint unter anderem auch deshalb nachteilig, weil berichtet wird, dass die Zielgruppe während der Programmlaufzeit in erheblicher Weise verändert wurde. Eine Einschätzung von Konzeptions- und Umsetzungsaspekten, die die Programmteilnahme und die Erreichung von Gesamtwirkungen in der FTI-Landschaft unter den unterschiedlichen Bedingungen beeinflusst haben können, hätte sich anhand einer Untersuchung von Nicht-Teilnehmern auf breitere Informationen stützen können. Es ist allerdings	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	zuzugestehen, dass eine vollständige Erfassung der ursprünglichen Zielgruppe anhand objektiver Daten vor großen Herausforderungen gestanden wäre. • Zur Bewältigung dieser Herausforderungen an die Analyse hätte z.B. internationale Fachliteratur einen Beitrag leisten können, die Aussagen über die Wirklogiken ähnlich gelagerter Programme in anderen Ländern trifft, oder über Verhaltensweisen der Zielgruppe an denen das Programm folgerichtig ansetzen kann. Dieser Schritt wurde nicht gesetzt. Auch internationale Interviews auf Basis einer Sampling-Strategie hätten ein probates Mittel darstellen können, um die Informationsbasis zu verbessern. • Die Analyse von Programmwirkungen erfolgt nicht unter Differenzierung nach Gesichtspunkten einer sich verändernden Programm-anlage (Zielsetzungen, Zielgruppe). Dadurch können Einfluss-faktoren, die die Erreichung von selbst gesetzten Programmzielen betreffen und für eine zukünftige Programmgestaltung relevant sein könnten, nicht erkannt worden sein. • Vorhandene Monitoringdaten, die für die Wirkungsanalyse herangezogen werden hätten können, werden von den EvaluatorsInnen unter Angabe von guten Gründen nicht genutzt (vgl. auch G3). Es wurden jedoch keine alternativen Daten erhoben, sodass vom Verzicht auf die Bezugnahme auf Monitoringdaten betroffene Betrachtungswinkel schlicht entfallen. Es werden teilweise Ergebnisse dargestellt, zu denen nicht klar nachvollziehbar ist und auch nicht berichtet wird, wie sie sich auf die Evaluationsschwerpunkte und Evaluationsfragestellungen beziehen. Die Evaluation konzentriert sich somit nicht durchgehend auf die möglichst gute Behandlung der wesentlichsten Gesichtspunkte. Zu einer am Berichtsanfang angekündigten Methode werden keine Ergebnisse berichtet.	
N5 Transparenz von Werten	Aus den vom Bericht angegebenen Evaluationsfragestellungen ist erkennbar, dass im Evaluationsauftrag mehrere Bewertungsmaßstäbe für einzelne Programmaspekte formuliert wurden. Weitere Evaluationsfragestellungen implizieren die Heranziehung eines Bewertungskriteriums, das jedoch in den Frageformulierungen nicht benannt ist. Vorab verankerte Kriterien, wie späterhin vorliegende Ergebnisse zu bewerten sein werden, werden nicht berichtet bzw. angewendet. Im Bericht wird nicht reflektiert, inwiefern diese verschiedenen Bewertungsmaßstäbe in einer Gesamteinschätzung des Programms auch zueinander konkurrieren können. Es wird nicht explizit gemacht, inwiefern im Ergebnis unterschiedliche Teilbewertungen nebeneinander stehen, und welchen Stellenwert sie für eine Gesamteinschätzung haben. Darüber hinaus wird in Bezug auf Programmwirkungen die Bewertbarkeit als solche als Evaluationsfrage ausgewiesen. Dies kann einerseits so aufgefasst werden, dass eine bewusste Auseinander-setzung mit der Frage nach adäquaten, unter Umständen alternativen Bewertungsmaßstäben vorhanden bzw. explizit gefordert war. Andererseits geht die Thematisierung von möglichen Bewertungsmaßstäben in der Evaluationsdurchführung an der Grundidee vorbei, dass die Wahl der Bewertungsmaßstäbe zu den Grundlagen einer Evaluation zählt, ebenso wie die eingesetzten Methoden. Die Frage nach den anzulegenden Bewertungsmaßstäben verleiht der vorliegenden Evaluation hier den Charakter eines ‚Evaluability Assessment‘.	3
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist klar strukturiert und bietet zentrale Ergebnisse in Form von Tabellen und Grafiken an. Tabellen- und Grafikverzeichnisse, anhand derer sich diese Ergebnisse rasch auffinden lassen, sind allerdings nicht beigegeben. Ein Executive Summary ist vorhanden. Der Bericht ist weitestgehend so formuliert, dass davon ausgegangen werden kann, dass er auch für Nicht-InsiderInnen gut verständlich ist. Einzelne Berichtspassagen sind allerdings so gehalten, dass nicht davon ausgegangen werden kann, dass	3

	AdressatInnen jenseits von AuftraggeberInnen und FTI-ExpertInnen gut folgen können. Dies betrifft auch die Erläuterung eines Sachverhalts, der aus Sicht der EvaluatorsInnen für die Einschätzung des Programms wesentlich ist, in den Schlussfolgerungen. Das Programm wird eingangs griffig dargestellt, sodass auch bislang nicht damit vertraute AdressatInnen der vorgelegten Analyse folgen können. Im weiteren Berichtsverlauf wird auf Veränderungen des Programms während seiner Laufzeit und auf dem Programm zugrunde liegenden Annahmen über seine Wirkungsweise eingegangen. Allerdings wird die genaue Verfasstheit des Programms in seinen unterschiedlichen Entwicklungsstadien hinsichtlich von vorhandenen bzw. nicht vorhandenen Gestaltungselementen und Relationen zwischen verfolgten Zielen und eingesetzten Mitteln nicht vollkommen greifbar.	
N8 Nutzung und Nutzen der Evaluation	Die vorhandenen Angaben zur Einbeziehung von Beteiligten und Betroffenen (N1) deuten darauf hin, dass sich die Evaluation auf Informationsbedürfnisse von zentralen Programmbeteiligten konzentriert hat. Darüber hinaus wurde im Zuge von Datenerhebungen ein Teil der vom Programm Betroffenen auf die Evaluation aufmerksam, sodass hier grundsätzlich auch ein Interesse an ihren Ergebnissen entstehen konnte. Es findet sich ein Hinweis, der darauf schließen lässt, dass eine laufende, nicht an formal definierte Zeitpunkte gebundene Interaktion mit Programmbeteiligten stattgefunden hat. Hinweise auf konkrete Vorstellungen über genauere Nutzungsweisen oder Schritte zu ihrer Förderung und Unterstützung finden sich nicht. Einige Einschränkungen der Nützlichkeit ergeben sich in Bezug auf mehrere Nützlichkeits-, Fairness- und Genauigkeitsstandards.	3
F3 Vollständige und faire Überprüfung	Der Bericht konzentriert sich auf die Identifikation von Schwächen des untersuchten Programms, was im Sinne einer kritischen Überprüfung als adäquat gelten kann. Er macht allerdings keine identifizierten Stärken des untersuchten Programms explizit. Der Bericht bringt Sichtweisen von Programmbeteiligten und von Betroffenen zur Geltung, wobei die Herkunft der jeweiligen Sichtweise transparent gemacht wird. Gelegentlich werden Ergebnisse in einer Weise dargestellt, aus der nicht gut erkennbar wird, welche Meinung sich die EvaluatorsInnen an dieser Stelle gebildet haben. Manche Ergebnisse werden lediglich deskriptiv dargestellt und es wird nicht klar, welche Gründe die EvaluatorsInnen haben, aus ihnen keine Schlüsse im Rahmen des angewendeten Analysekonzepts zu ziehen. An mehreren Stellen wird nicht klar dargestellt, welchen genauen Stellenwert bestimmte Beobachtungen haben. Bei qualitativen Auswertungen werden Zitate gebracht, es wird aber nicht angegeben, wie typisch diese Erfahrungen und Sichtweisen sind bzw. wie häufig oder selten sie anzutreffen waren. Aspekte der Programmumsetzung werden teilweise in unscharfen Formulierungen dargestellt, aus denen nicht klar wird, ob es sich um eine durchgängige oder partielle Praxis gehandelt hat.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Vier der dargestellten Evaluationsfragestellungen betreffen die Kontextualisierung des Programms durch benachbarte Programme. Sie machen in der formulierten Form zu ihrer guten und gesicherten Beantwortbarkeit im Grunde eine Untersuchung erforderlich, die sich mit beobachtbaren Wirkungen und dahinter stehenden Wirklogiken der benachbarten Programme intensiv auseinandersetzt, im Optimalfall genauso intensiv wie mit dem untersuchten Programm selbst. Eine derartige Untersuchung wurde nicht angestellt.	5
G3 Beschreibung von Zwecken und Vorgehen	In einem eigenen Teilkapitel wird die Vorgehensweise der	3

	Evaluation dargestellt. Dabei handelt es sich einerseits um eine Darstellung einer Methode im engeren Sinne, und andererseits um eine Listung von herangezogenen Unterlagen und weiteren gesetzten Erhebungsschritten, die eher kursorisch beschrieben werden. Eine noch klarere und leicht auffassbare Darstellung wäre möglich gewesen. Gemeinsam mit den im Anhang beigegebenen Unterlagen wird das Vorgehen der Evaluation grundsätzlich nachvollziehbar, und in den meisten – wenn auch nicht allen – Hinsichten im Detail. Weitere methodische Erläuterungen werden im Berichtsverlauf zu einzelnen Erhebungen bzw. Datentypen gegeben. Ein klar erkennbarer Bezug zwischen Evaluationsschwerpunkte, Evaluationsfragestellungen und Methoden, die die Beantwortung bzw. Einschätzung ermöglichen sollen, wird nicht hergestellt. Die Tragfähigkeit einzelner Informationsquellen wird eingeschätzt. Es wird unter anderem begründet, warum manche verfügbare Monitoringdaten für die Analyse nicht genutzt werden. Stärken und Schwächen des Methodeneinsatzes als Gesamtheit werden nicht diskutiert. Trotz der verschiedenen vorhandenen Angaben und punktuellen Erläuterungen wird es den BerichtsadressatInnen letztlich nicht gut und vollständig ermöglicht, etwaige Vor- oder Nachteile aller einzelnen Methoden und des gewählten Methodensets für die Behandlung der übergreifenden Evaluationsschwerpunkte und der einzelnen Evaluationsfragestellungen einzuschätzen. Weiterführende Angaben auch über Eckpunkte des Evaluationsprozesses wären im vorliegenden Fall wünschenswert gewesen, um die Evaluation für nicht mit ihrem Vorgehen vertraute Personen besser einschätzbar zu machen. Es lässt sich abgesehen vom Vermerken des konkreten Auftraggebers nicht erkennen, inwiefern alle Programmeigentümer und alle relevanten Akteure der Ausgestaltung und Umsetzung des Programms an der Gestaltung des Auftrags beteiligt und in die Untersuchung des Programms einbezogen waren. Das Vorgehen der Evaluation ist unentschieden im Bezug auf die Veränderungen von Zielsetzungen und Zielgruppen des Programms während seiner Laufzeit (vgl. auch N4). Sie bietet keine ausdrückliche Darlegung ihres Umgangs mit dieser Herausforderung und Weichenstellung für die Analyse an, und mit etwaigen Konsequenzen der gewählten Vorgehensweise. Die Unentschiedenheit des Vorgehens kann Auswirkungen auf die Einordnung von Beobachtungen hinsichtlich von Vor- und Nachteilen von Gestaltungsweisen des Programms gehabt haben, aus denen gemäß der angegebenen Evaluationszwecke für die Zukunft gelernt werden kann.	
G8 Begründete Schlussfolgerungen	Zwei abschließende Kapitel widmen sich der Diskussion der Gesamtergebnisse und Schlussfolgerungen, klar getrennt von den vorangegangenen Datendarstellungen und –interpretationen. Die Schlussfolgerungen sind intern nach Analyseschwerpunkten gegliedert. Eine direkte Korrespondenz zu den ausgewiesenen Evaluationsfragestellungen wird nicht hergestellt. Es finden sich nahezu durchgehend Schlussfolgerungen, die aus dem vorher Berichteten heraus gut begründet sind. Allerdings taucht auch eine Stellungnahme zu einer einzelnen Programmlinie auf, die nicht Gegenstand einer ausgewiesenen Evaluationsfrage war und zu der nun ausgesagt wird, ihre Interventionslogik bleibe zu überprüfen. Dies kann wohl als Hinweis auf in früheren Phasen der Evaluation doch nicht vollkommen geklärte Informationsbedürfnisse und Arbeitsaufträge verstanden werden (vgl. N1). Zu einzelnen Punkten werden bislang noch nicht berichtete Fakten herangezogen oder Hintergrundwissen der EvaluatorenInnen über nicht explizit untersuchte Aspekte des FTI-Systems eingebracht. Die ausgesprochenen Empfehlungen konzentrieren sich auf die Relevanz des Programms. Sie beziehen sich im Grunde auf fünf von allen Evaluationsfragestellungen, die sie teilweise anhand prinzipieller Überlegungen einschätzen, aber nicht anhand in der Untersuchung erbrachter Daten. Sie geben keine spezifischen, umsetzbaren Hinweise für den Fall, dass die Auftraggeber das	3

	untersuchte Programm in optimierter Form weiterführen möchten.	
--	--	--

Evaluationsbericht Nr. 7		Beobachtungszeitraum: 2011 - 2014
Klassifikation	Einordnung	
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)	
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact	
Standard	Kommentar	Einstufung*
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen von Datenerhebungsverfahren einbezogen wurden Programmverantwortliche und -umsetzende, unterschiedliche Zielgruppen, Begünstigte (Fördernehmer), sowie einige FTI-politische Akteure. Unter Programmverantwortlichen und –umsetzenden wurden auch Akteure einbezogen, die in früheren Phasen des untersuchten Programms tätig waren. Die vorfindlichen Angaben reichen jedoch nicht aus, um einzuschätzen, ob alle relevanten Akteure vollständig einbezogen wurden oder eine selektive Auswahl erfolgte. Die Einbeziehung von Zielgruppen kann als selektiv eingeschätzt werden. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure oder andere politische Akteure einbezogen worden wären, die nicht unmittelbarer mit dem Programm befasst waren, aber – auch wegen der übergreifenden Thematik des Programms - an der Evaluation und ihren Ergebnissen ein Interesse haben könnten.	3
N2 Klärung der Evaluationszwecke	Angegeben wird: Bewertung der mittelfristigen Effekte und Wirkungsrichtungen, um einen Beitrag für die Weiterentwicklung des Programms zu leisten. Analyse der bisherigen Maßnahmen auf unterschiedlichen Ebenen hinsichtlich des Leistungs- und Produktportfolios, sowie Feststellung des Nutzungsverhaltens und der Veränderungsbereitschaft der direkten Zielgruppen. Die Angabe ist als global und unscharf zu bezeichnen und eine Priorisierung innerhalb der Bandbreite der auf unterschiedlichen Ebenen angegebenen Themensetzungen ist nicht zu erkennen. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären.	4
N4 Auswahl und Umfang der Informationen	Die Herangehensweise der Evaluation wird abgesehen von der eher globalen Zwecksetzung (vgl. N2) und der Darstellung der Methodenwahl nicht näher beschrieben. Evaluationsschwerpunkte in Bezug auf das komplexe, aus mehreren Maßnahmen für mehrere Zielgruppen aufgebaute Programm oder Evaluationsfragestellungen werden nicht dargestellt. Die Evaluation setzt elf empirische Methoden parallel ein. Der Mix aus quantitativen und qualitativen Methoden erscheint auch in seiner Breite in Bezug auf die umrissene Zwecksetzung der Evaluation und die dargestellte Programmanlage grundsätzlich gut geeignet. Die Vorgehensweise der Evaluation ist in einem ersten Abschnitt dadurch gekennzeichnet, dass sie sich mit den dem Programm zugrunde liegenden Annahmen, mithilfe welcher Mittel welche Zielsetzungen erreicht werden sollen bzw. können, und mit konzeptuellen Grundlagen im internationalen Rahmen auseinandersetzt. Die Zielsetzungen des Programms werden systematisch in übergeordnete, intermediäre und operative Ziele gegliedert. Direkte und indirekte Zielgruppen werden gegliedert. In einer Logic Chart werden hierarchisierte Zielsetzungen in Relation zu den Programmlinien gesetzt und Programmaktivitäten mit ihren erwarteten Outputs und Outcomes zugeordnet. In späteren Berichtsabschnitten werden Logic Charts differenziert zu einzelnen	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Programmlinien präsentiert. Im Rahmen dieser Vorgehensweise werden auch Unklarheiten in der vorfindlichen Programmformulierung herausgearbeitet. Die empirische Untersuchung in den folgenden Berichtsteilen löst die Möglichkeiten, die sich aus der strukturierten Aufschlüsselung der Wirklogik(en) des Programms ergeben, und die die EvaluatorsInnen ihr auch explizit zuschreiben, allerdings kaum ein. Es ist nicht erkennbar, dass sie zum Ausgangspunkt einer gezielten Wahl von Analysepunkten und eines darauf bezogenen Methodendesigns geworden wäre. Große Teile des Berichts begnügen sich mit einer Darstellung der Programm-Outputs. Auch Programmlinien, die nicht Förderungsvergabe beinhalten, werden zwar in dieser Evaluation eingehend untersucht, aber doch ebenso in erster Linie hinsichtlich ihrer Outputs. Auf Outcomes und erste Wirkungen wird nur unzureichend differenziert eingegangen. Getrennte Ergebnisse zu Outputs, Outcomes und beobachtbaren Wirkungen werden kaum integriert bzw. aufeinander bezogen. Die erklärte Absicht, tiefer liegende Wirkmechanismen zu erklären und das gesamthafte Zusammenwirken der Programmbestandteile zu erkennen, kann so nicht eingelöst werden. Möglichkeiten zur Gewinnung weiterer aussagekräftiger Daten, die im Rahmen der eingesetzten Methoden grundsätzlich bestanden hätten, wurden nicht genutzt. Es wird unter anderem festgehalten, dass in internationalen Studien eine Vielzahl von Hinweisen über Wirkmechanismen im Themenbereich erbracht wurden. Die vorliegende Bericht nimmt zwar an verschiedenen Stellen immer wieder Bezug auf solche Mechanismen, eine systematische Aufarbeitung ist allerdings nicht erfolgt. Hierzu wäre eine spezifische, zusätzliche Methode notwendig gewesen. Es werden unter anderem Outputs an mehreren Stellen nach Verteilungen untersucht, die im Wirkungslogik-Konzept bzw. den angegebenen Evaluationszwecken keine klare Motivation haben. Die Evaluation konzentriert sich somit nicht durchgehend auf die möglichst gute Behandlung der wesentlichsten Gesichtspunkte.	
N5 Transparenz von Werten	Ausgehend von der ausgewiesenen Evaluationsaufgabe, das Programm zu bewerten, verfährt die Berichterstattung ohne explizite Darlegungen des Vorgehens in gleichsam selbstverständlicher Weise. Einige Berichtsabschnitte sind von den Sichtweisen der EvaluatorsInnen getragen, andere bringen die Sichtweisen von Betroffenen (FördernehmerInnen, Zielgruppen) und Programmverantwortlichen zum Ausdruck. In diesen Fällen wird die Herkunft der jeweiligen Betrachtungsweise ausgewiesen. Hinsichtlich der Konzeptualisierung des Programms bringen die EvaluatorsInnen eine abgesicherte und anerkannte Begriffssystematik zum Einsatz. Der Bericht legt allerdings nicht klar dar, inwiefern die angelegte Systematik die Programmanlage selbst widerspiegelt, und inwiefern sie zu Analyse Zwecken ex post von außen angelegt wird. Einzelne Berichtsabschnitte sind an der Erreichung von Programmzielen ausgerichtet, ohne dass allerdings Zielerreichung als tragende systematische Herangehensweise dargestellt bzw. klar erkennbar würde. Im Berichtsverlauf werden unterschiedliche Bewertungsmaßstäbe für verschiedene Einzelbeobachtungen herangezogen, die dem vorgestellten Untersuchungskonzept, die Wirklogik des Programms zu beleuchten, oft schlecht entsprechen (z.B. kann Zufriedenheit von TeilnehmerInnen keine Erkenntnisse erbringen, inwiefern die Teilnahme zum Eintreten von Bewusstseinsständen oder Verhaltensweisen beiträgt). In einzelnen Berichtsabschnitten bleibt unklar, welche Maßstäbe zur Einschätzung der Ergebnisse herangezogen werden, und zu welcher Einschätzung die EvaluatorsInnen gelangen. Eine Auseinandersetzung damit, wie sich die verschiedenen Einzelbewertungen zu einer Gesamteinschätzung des Programms fügen, wird nicht angeboten.	4

N6 Vollständigkeit und Klarheit der Berichterstattung	Es liegt ein überarbeiteter Bericht vor. Gründe für die Überarbeitung und etwaige Veränderungen gegenüber dem Originalbericht werden nicht angegeben. Es kann nicht beurteilt werden, ob es sich nur um Zusätze zum Originalbericht handelt oder ob auch Eingriffe in dessen ursprüngliche Darstellungen erfolgt sind. Ein Executive Summary ist beigegeben. Der Bericht ist nach Abschnitten der Programmlogik und Programmbestandteilen gegliedert, was die Auffassung der Berichtsinhalte und Analyseschritte für AdressatInnen zweifellos erleichtert. Das Programm wird dargestellt, sodass auch mit ihm bislang nicht Vertraute der Untersuchung folgen können. Tabellen und Grafiken werden eingesetzt, um zentrale Inhalte und Ergebnisse leicht auffassbar zu vermitteln. In manchen Berichtsabschnitten wurde diese Präsentationsstrategie allerdings nicht durchgehalten, zusätzliche Tabellen und Grafiken wären auch hier für eine gute Auffassbarkeit der textlich referierten Ergebnisse jedenfalls dienlich gewesen. Es finden sich auch vereinzelt Grafiken, die nicht ausreichend beschriftet sind. In den Formulierungsweisen des Berichts wird des Öfteren nicht klar zwischen prinzipiellen Überlegungen auf Konzeptebene und empirischen Beobachtungen bzw. deren Interpretation differenziert. Dies kann dazu führen, dass BerichtsadressatInnen Mühe haben, die Argumentationsweise zu verstehen (vgl. auch N5 und G3).	3
N8 Nutzung und Nutzen der Evaluation	Aus den Angaben zur Einbeziehung von Beteiligten und Betroffenen (N1) kann geschlossen werden, dass breite Kreise auf die Evaluation aufmerksam gemacht wurden, sodass sie in der Folge grundsätzlich ein Interesse an ihren Ergebnissen haben konnten. Eine eingehende Auseinandersetzung mit der Programmlogik gemeinsam mit Programmverantwortlichen und –umsetzenden bietet prinzipiell gute Möglichkeiten, die Evaluation für sie nützlich zu machen. Angaben, dass hier lernorientierte Interaktionen stattgefunden haben (z.B. Workshops), finden sich allerdings nicht. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden soll, werden nicht dargestellt. Es liegt ein Bericht vor, der in Bezug auf mehrere Nützlichkeits-, Fairness- und Genauigkeitsstandards einige Schwächen aufweist.	4
F3 Vollständige und faire Überprüfung	Der Bericht ist grundsätzlich in einer neutralen Sprache gehalten. Er benennt primär Schwächen des Programms, was grundsätzlich als Entsprechung zu einem lernorientierten Ansatz gesehen werden kann. Sichtweisen von erreichten und nicht erreichten Teilen von verschiedenen Zielgruppen sowie von Programmverantwortlichen werden dargestellt und mit klaren Hinweisen auf ihre Herkunft versehen. Zu manchen berichteten Datenlagen nehmen die EvaluatorsInnen keine Einschätzung vor. Die Ergebnisse einiger Untersuchungsschritte werden nicht vollständig dargestellt. Die neutral formulierende und auf Fakten verweisende Art der Berichterstattung wird an einzelnen Stellen durchbrochen, an denen wertende Feststellungen erfolgen, ohne dass die dahinter stehenden Fakten erkennbar würden, sei es durch deren unmittelbare Darstellung oder durch einen Verweis. Manche Berichtsstellen kommentieren Aspekte der Programmgestaltung in positiver Weise, an anderen wird zu eher negativen Formulierungsweisen tendiert.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das Programm wird in seiner Motivation abrisshaft in den Kontext internationaler Entwicklungen und identifizierter Herausforderungen gestellt. Drei internationale Fallbeispiele aus verschiedenen Ländern werden herangezogen, um die Programmlogik in einen internationalen	3

	Rahmen einzuordnen. Die hier erbrachten Informationen reichen aus, um einen Vergleich auf der Ebene grundsätzlich verfolgter Strategien anzustellen, jedoch nicht, um die genaueren Wirklogiken der Vergleichsprogramme und die Rolle von Faktoren zu erfassen, die deren Erfolg jeweils fördern oder behindern. Dazu hätte es wesentlich umfangreicherer Daten und Analysen bedurft. Das Programm soll von seiner Anlage her Bewusstseins- und Verhaltensänderungen in mehreren Zielgruppen auslösen. Es kann berechtigt angenommen werden, dass zahlreiche andere, vom Programm nicht erreichte Prozesse und Strukturbedingungen die Denk- und Verhaltensweisen dieser Zielgruppen mit bestimmen, sodass sie für die Passgenauigkeit der Programmgestaltung und die Chancen für eine gute Zielerreichung Bedeutung haben. Entsprechende Untersuchungen, aus denen solche Faktoren erkannt werden könnten, wurden ein Stück weit angestellt, hätten aber deutlich umfassender sein können. Bedingungen, die für verschiedene Zielgruppen gelten, werden nur stellenweise angedeutet.	
G3 Beschreibung von Zwecken und Vorgehen	Der Evaluationsansatz wird zunächst durch eine Übersichtsgrafik greifbar, die die verschiedenen eingesetzten Methoden in Relation zu den Untersuchungsschwerpunkten darstellt. Methodische Angaben finden sich auch im Anhang sowie in einzelnen Berichtsabschnitten, wo sie nicht immer leicht auffindbar sind. Methodische Angaben sind in einzelnen Berichtsabschnitten und im Anhang vorhanden. Sie sind jedoch nicht für alle eingesetzten Methoden ausreichend, um die Tragfähigkeit aller berichteter Ergebnisse einschätzen zu können. Bei der Darstellung von einzelnen Ergebnissen wird nicht immer klar vermittelt, was genau untersucht wurde (z.B. exakte Fragestellungen hinter Daten, über die in unscharfer Diktion berichtet wird). Die Evaluation nimmt eine Strukturierung der Programmanlage vor (vgl. N4). Dem wird allerdings in den verschiedenen Methodendarstellungen an verschiedenen Orten des Berichts kein systematischer Platz in der angestellten Untersuchung zugewiesen. Bei der Strukturierung der Programmanlage wird auch nicht deutlich gemacht, inwieweit die der Untersuchung zugrunde gelegte Programmstruktur aus Programmunterlagen bezogen, im Zuge der Evaluation gemeinsam mit Programmverantwortlichen herausgearbeitet, oder von den EvaluatorenInnen supponiert bzw. aus Beobachtungen des Programmgeschehens destilliert wurde. An unterschiedlichen Berichtsstellen finden sich dazu nicht völlig konsistente Angaben. Ein Teil der Untersuchung ist auf eine Konsistenzprüfung der Programmanlage ausgerichtet, und es wird berichtet, dass sich das Programm während seiner Laufzeit verändert hat. Es wird jedoch nicht berichtet, welche Konsequenzen die Evaluation aus diesen Feststellungen für ihre analytische Vorgehensweise gezogen hat (z.B. Wahl bestimmter Klassifikationsraster für Programmereignisse). In der Berichterstattung wird nicht so klar differenziert, was für das Programm hinderlich war bzw. gewesen sein konnte, und was für die Durchführung der Evaluation hinderlich war, dass alle BerichtsadressatInnen mit Sicherheit gut folgen können. In der Berichterstattung über Outcomes weisen die EvaluatorenInnen zurecht darauf hin, dass die kausale Zuordnung von beobachteten Veränderungen zum Programm nur eingeschränkt möglich ist. In manchen Berichtsteilen werden Aussagen zur Nicht-Verfügbarkeit von Daten getroffen, doch ist dies nicht überall der Fall. Es wird nicht klar vermittelt, wie vollständig und aussagekräftig die verschiedenen Programmbestandteile tatsächlich untersucht werden konnten, und wie umfassend oder eingeschränkt die Untersuchung des Gesamtprogramms daher ist. Es kann davon ausgegangen werden, dass die Vorgehensweise der Evaluation von BerichtsadressatInnen nicht leicht aufgefasst werden kann und deren Vor- und etwaige Nachteile durch sie nur schlecht eingeschätzt werden können.	4
G8 Begründete Schlussfolgerungen	Schlussfolgerungen stehen zum Einen am Ende von Einzelabschnitten der Analyse, wo sie als solche auch ausgewiesen und leicht auffindbar sind. Die EvaluatorenInnen stützen sich hier oftmals	3

	auf Verbesserungsvorschläge aus den Zielgruppen, was sie auch klar ausweisen. Teilweise stützen sich Schlussfolgerungen auch auf Annahmen der EvaluatorInnen, ohne dass ein Bezug auf berichtete Daten transparent gemacht wird. Zum Anderen widmet sich ein abschließendes, von den Ergebnisdarstellungen getrenntes Kapitel Schlussfolgerungen und Empfehlungen. Hier wird übersichtlich nach identifizierten Stärken und Schwächen sowie Empfehlungen gegliedert. Viele, aber nicht alle Schlussfolgerungen sind hier hinsichtlich zugrunde liegender Fakten transparent. Bei aller Plausibilität, die <i>prima vista</i> vorliegt, werden doch teilweise bisher so noch nicht berichtete Sachverhalte verhandelt, und teilweise könnten Überinterpretationen von punktuellen Beobachtungen vorliegen (vgl. N4, G2). Der Konnex zur präsentierten Wirklogik ist eher schwach ausgeprägt. Die Empfehlungen beinhalten vor allem allgemein-prinzipielle Überlegungen zur Programmanlage, daneben werden auch einige spezifische Hinweise zu Gestaltungsweisen einzelner Programmlinien gegeben. Dabei wird auch in allgemein-unscharfer Weise an nicht direkt mit dem Programm befasste Akteursgruppen appelliert, die in die Evaluation nicht einbezogen waren.	
--	--	--

Evaluationsbericht Nr. 8		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In Datenerhebungsverfahren wurden Programmverantwortliche und mit seiner Umsetzung Betraute, Förderempfänger und nicht erfolgreiche Antragsteller, sowie mit dem Programm befasste externe ExpertInnen einbezogen. Es wird berichtet, dass eine Reflexion von Zwischenergebnissen gemeinsam mit für Programmanlage und –umsetzung Verantwortlichen im Lauf des Evaluationsprozesses erfolgte. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben könnten.	3	
N2 Klärung der Evaluationszwecke	Angegeben wird: Den bisherigen Programmverlauf reflektieren und Schlussfolgerungen und Empfehlungen für die Weiterentwicklung des Programms erbringen. Eine Priorisierung von Untersuchungsschwerpunkten innerhalb der sehr breit formulierten Aufgabenstellung ist nicht zu entnehmen. Es liegt kein Hinweis vor, dass in eine Klärung der Evaluationszwecke neben den direkten AuftraggeberInnen auch andere Beteiligte und Betroffene des untersuchten Programms einbezogen worden wären.	4	
N4 Auswahl und Umfang der Informationen	Der Evaluationsauftrag geht auf eine Vielzahl von Evaluationsfragestellungen zurück, die die EvaluatorInnen in Themenblöcke gliedern, um eine gut strukturierte Analyse anstellen zu können. In dieser Gliederung werden einige, aber nicht alle vorgesehenen präzisen Detailfragestellungen berichtet. Es werden mehrere Methoden parallel eingesetzt, die mit ihrem Mix aus qualitativen und quantitativen Ansätzen gut geeignet erscheinen. Die Evaluation ist sichtlich bestrebt, sich der vorhandenen Komplexität des untersuchten Programms zu stellen und sie ein Stück weit bewältigbar zu machen. Das Programm wird in Motivation und Aufbau dargestellt. Ergebnisse früherer Evaluationen von Vorgängerprogrammen werden dabei aufgegriffen. Zusätzlich wurden die ehemaligen EvaluatorInnen in die Vorgehensweise der Evaluation einbezogen. Zielsetzungen und Umsetzungsstrukturen des untersuchten Programms werden detailliert aufgeschlüsselt. Die Evaluation setzt sich mit Wirkmechanismen und intervenierenden Faktoren (v.a. Bedingungen für die Erreichung mittel- und längerfristiger Wirkungen) anhand adäquater Daten auseinander. Wesentliche Aspekte des Umsetzungsprozesses – wenn auch nicht alle prinzipiell denkbaren - wurden mithilfe adäquater Methoden untersucht. Unter anderem werden auch Unklarheiten in der Programmkonzeption herausgearbeitet. Um die Wirkungsweise des Programms möglichst gut zu verstehen, hätte unter dem Betrachtungswinkel von Wirkungen noch eingehender zwischen Outputs, Outcomes unterschiedlicher logisch-hierarchischer Stellung, und Impacts differenziert werden können. Während ein adäquater Umgang damit, inwiefern beobachtete Effekte tatsächlich dem untersuchten Programm zugeschrieben werden können, grundsätzlich vorliegt, hätte an manchen Stellen noch differenzierter mit dieser Kausalitätsproblematik umgegangen werden können. In der Aufschlüsselung des Programmgeschehens werden unter anderem auch Fragestellungen untersucht, zu denen nicht deutlich wird, welchen Stellenwert sie in einem Analysekonzept haben, das auf die Funktionsweise des Programms, ihm zuschreibbare Wirkungen, und das Erkennen von etwaigen förderlichen und	2	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	hinderlichen Faktoren bei der Verfolgung von Zielen abzielt. Die Evaluation konzentriert sich somit nicht durchgehend auf die möglichst gute Behandlung der wesentlichsten Gesichtspunkte.	
N5 Transparenz von Werten	Die Evaluation stützt sich auf objektive Daten ebenso wie auf Erfahrungen und Sichtweisen von Programmverantwortlichen, Zielgruppen und externen Expertinnen, die hinsichtlich ihrer Herkunft transparent gemacht werden. Der Bericht setzt sich mit Bewertungsmaßstäben auseinander, die innerhalb des untersuchten Programms zur Anwendung gebracht wurden. Die Vorgehensweise der Evaluation ist dadurch gekennzeichnet, dass die Art und Weise, wie das Programm mithilfe der eingesetzten Mittel seine selbstgesetzten Ziele erreicht, differenziert analysiert werden soll. Dabei werden zwangsläufig Einordnungen verschiedener beobachteter Sachverhalte in eine übergreifende Bewertungssystematik notwendig. An manchen Stellen werden Abwägungen über adäquate Bewertungsmaßstäbe und Entscheidungen dargelegt. Andere, zum Teil in FTI-Evaluationen „übliche“ Bewertungsmaßstäbe werden kommentarlos angewendet. Eine wesentliche Rolle für die Schlussfolgerungen spielt die Frage einer Abgrenzung des Programms von anderen Förderungen. Es wird jedoch nicht vermittelt, ob dieses Kriterium Bestandteil der ursprünglichen Programmanlage ist oder zu Analysezielen von außen an das Programm herangetragen wurde. Es unterbleibt eine Darstellung, in der den BerichtsadressatInnen für alle Schritte der Evaluation klar aufgeschlüsselt würde, wie die Gesamtanalyse konfiguriert ist und welche Auswirkungen dies auf die erbrachten Ergebnisse, Schlussfolgerungen und Empfehlungen gehabt haben könnte.	3
N6 Vollständigkeit und Klarheit der Berichterstattung	Die Berichtssprache ist klar und allgemein verständlich. Grafiken und Tabellen werden zur Vermittlung von Ergebnissen eingesetzt. Grafiken- und Tabellenverzeichnisse, anhand derer sie rasch aufgefunden werden können, wurden allerdings nicht beigegeben. Einzelne Tabellen und Grafiken hätten eine bessere Beschriftung verdient. Ein Executive Summary ist nicht beigegeben. In einer klaren Berichtsstruktur behandeln einzelne Berichtsabschnitte die Programmanlage, die Programmumsetzung, sowie Ergebnisse und Wirkungen, die beobachtet bzw. dem Programm zugeschrieben werden können. Daran schließt sich eine integrierende Betrachtung der Gesamtergebnisse anhand übergreifender Fragestellungen. Es erfolgt eine Darstellung des untersuchten Programms, anhand derer sich auch bislang mit dem Programm nicht Vertraute leicht ein gutes Bild verschaffen können, das ihnen das Auffassen der Evaluation wesentlich erleichtert. Zielsetzungen des Programms werden gut nachvollziehbar aus Programmdokumenten heraus benannt, Anpassungen des Programms während seiner Laufzeit sind integriert. Die Tatsache, dass offenbar an verschiedenen Stellen implizit verbleibende Fragestellungen verhandelt werden, verleiht dem Bericht allerdings auch ein Stück weit den Charakter eines Insider-Berichts an die AuftraggeberInnen, der nur von diesen vollständig aufgefasst werden kann. Auch für sie dürften die Beantwortungen einzelner Fragestellungen nur durch ein eingehendes Berichtsstudium identifizierbar werden, und es kann davon ausgegangen werden, dass Zusatzkommunikation über die Vorlage des vorliegenden Berichts hinaus hier jedenfalls sinnvoll gewesen wäre (hierzu finden sich keine Angaben).	2
N8 Nutzung und Nutzen der Evaluation	Es liegt ein Bericht vor, der für die HauptadressatInnen (AuftraggeberInnen) und weitere Akteure im Umfeld des Programms jedenfalls nützlich sein kann. Aus den Angaben zur Einbeziehung von beteiligten und Betroffenen (N1) kann geschlossen werden, dass die HauptadressatInnen gut eingebunden wurden und darüber hinaus auch einige Akteursgruppen auf die Evaluation aufmerksam gemacht wurden, wodurch Interesse an ihren Ergebnissen geweckt werden konnte. Es liegt allerdings kein Hinweis vor, dass breitere Gruppen	2

	potenzieller NutzerInnen im Bereich der der FTI-Politik, die auch über benachbarte Programme zu bestimmen haben und nicht mit den einbezogenen Personen identisch sind, einbezogen oder mit dem Bericht adressiert wurden. Hinweise auf konkrete Vorstellungen über genauere Nutzungsweisen oder Schritte zu ihrer Förderung und Unterstützung finden sich nicht. Leichte Einschränkungen ergeben sich vor allem aus den Erfüllungsgraden der Standards G2 und G3.	
F3 Vollständige und faire Überprüfung	Der Bericht erfolgt in neutralem Tonfall und trennt Fakten von Bewertungen. Er identifiziert sowohl Stärken als auch Schwächen des Programms, wobei die Identifikation von Schwächen, aus denen für die Zukunft gelernt werden kann, im Vordergrund steht. Schwächen des Programms werden auch anhand von Sichtweisen von Beteiligten und Betroffenen herausgearbeitet, wobei die Herkunft von berichteten Sichtweisen transparent gemacht wird. Die Interpretation qualitativer Daten ist von einer sachdienlichen Strategie getragen, vernachlässigt aber Hinweise auf das Gewicht der beschriebenen Effekte und Faktoren im Gesamtbild aller erhaltenen Angaben. Nicht zu allen Ergebnissen drücken die EvaluatorsInnen klar aus, wie sie diese Ergebnisse einschätzen und welche Schlüsse sie daraus ziehen. Auch Schlussfolgerungen sind gelegentlich in einer metaphorischen Sprache gehalten, die darauf zu vertrauen scheint, dass die HauptadressatInnen des Berichts schon verstehen werden, was gemeint ist. Dies geht auf Kosten einer verlässlichen Erkennbarkeit der Berichtsaussagen bei allen BerichtsempfängerInnen. Die Identifikation mancher Schwächen erscheint verbunden mit implizit verbleibenden Konzeptfragen. Manches hätte hier unter Umständen auch im Sinne positiver Nebenwirkungen aufgefasst werden können. Einer Identifikation von Nebenwirkungen wird freilich in der gesamten Analysestrategie, so wie sie präsentiert bzw. erkennbar ist, kein besonderer Stellenwert zugemessen.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor.	3
G2 Kontextanalyse	Durch die gewählte Vorgehensweise werden auch essentielle Randbedingungen greifbar, unter denen das Programm in seinen Zielgruppen Wirkungen erreichen kann bzw. daran gehindert ist. Es wird im Zuge der BerichtsDarstellungen deutlich, dass das untersuchte Programm durch potenziell überlappende Programme kontextualisiert ist, und dass die Entfaltung von angestrebten Programmwirkungen mit diesen anderen Programmen auch verknüpft sein kann. Die Stellung im Portfolio wird auf einer allgemein-prinzipiellen Ebene und auf der Basis von Hintergrundwissen der EvaluatorsInnen angesprochen. Eine systematische Aufarbeitung der faktischen Positionierung und der spezifischen, evident gemachten Wirkungsweise des Programms im Förderportfolio erfolgt nicht. Dies obwohl ausgewiesene Evaluationsfragestellungen in die Richtung solcher Analysen weisen. Eine Überschneidung mit einem internationalen Projekt wird thematisiert. Dabei wird allerdings nicht klar, wie dieses Projekt in systematischer Hinsicht im Verhältnis zum untersuchten Programm begriffen wird.	3
G3 Beschreibung von Zwecken und Vorgehen	Es wird berichtet, dass der Evaluation eine große Zahl an Evaluationsfragestellungen zugrunde liegen. Diese werden hinsichtlich übergeordneter Themensetzungen skizziert, aber nicht vollständig präzise angegeben. Die EvaluatorsInnen haben offenbar eine Strukturierung der vielfältigen Evaluationsfragestellungen vorgenommen, um die Untersuchung besser strukturieren zu können. Das resultierende Analysekonzept wird als solches greifbar gemacht, der dahinter stehende Auftrag jedoch nur in Ansätzen. Ein kurzes Methodenkapitel gibt eine klare und übersichtliche Darstellung der gesetzten Erhebungsschritte mit Überlegungen zu	3

	ihrer Motivation. Methoden hier und an weiteren Berichtsstellen soweit dargestellt, dass ihre Tragfähigkeit weitestgehend nachvollziehbar werden. Einige zusätzliche Angaben wären jedoch wünschenswert gewesen, um die Vorgehensweise der Evaluation vollkommen transparent und einschätzbar zu machen (z.B. Sampling-Strategien für mehrere eingesetzte Methoden). Eine explizite Reflexion von Vor- und etwaigen Nachteilen des eingesetzten Methodensets und etwaiger Folgen für die erbrachten Ergebnisse und Einschätzungen wird nicht angeboten.	
G8 Begründete Schlussfolgerungen	Die integrative Betrachtung der Gesamtergebnisse ist nach drei übergeordneten Fragestellungen strukturiert, die allerdings die eingangs vorgestellte Strukturierung des Untersuchungsaufbaus nur bedingt abbilden. Es wird angegeben, dass die Gesamtbetrachtung an den Evaluationsfragestellungen orientiert ist, doch verbleibt dies implizit und ist somit nur für die AuftraggeberInnen erkennbar. Hier vorfindliche Aussagen haben einerseits guten Rückhalt in zuvor berichteten Daten, andererseits werden neue Fakten und Überlegungen eingebracht. Der Gesamtgehalt der zuvor dargelegten Analysen von Programmaspekten als Wirklogiken wird nicht vollständig transportiert. Ein eigenes Kapitel widmet sich zentralen Schlussfolgerungen und Empfehlungen. Diese befassen sich in erster Linie mit der Programmanlage, wobei der Bezug zu zuvor berichteten Fakten kaum transparent gemacht wird. Sie beziehen sich stark auf die Kontextualisierung des Programms durch andere FTI-Förderungen, die allerdings nur wenig untersucht wurde, und setzen wie selbstverständlich ein Argument ein, von dem nicht ganz klar wird, woher es stammt und in welchem Verhältnis es zur übergeordneten Zielsetzung des Programms steht (vgl. N5). Daneben werden einige umsetzbare Empfehlungen zu Umsetzungsaspekten gegeben. Eine volle Bezugnahme auf die eingangs vorgestellten Untersuchungsschwerpunkte bzw. Fragestellungen ist nur eingeschränkt gegeben.	3

Evaluationsbericht Nr. 9		Beobachtungszeitraum: 2007 - 2010	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen von Datenerhebungsverfahren wurden Programmverantwortliche und –umsetzende und vom Programm Betroffene (Fördernehmer und Zielgruppen) einbezogen. Des Weiteren wurden verschiedene FTI-politische Akteure einbezogen, die in der Entstehungsphase des Programms eine Rolle gespielt haben, oder aktuell eine Interesse am Programm haben konnten. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in die Planung der Evaluation hinsichtlich von Informationsbedürfnissen einbezogen worden wären.	2	
N2 Klärung der Evaluationszwecke	Angegeben wird: Analyse der Konzeption und Prozesse des Programms sowie Unterstützung der strategischen Steuerung und förderungspolitischen Positionierung des Programms. Handlungsleitende Perspektiven und Gestaltungsinformationen bezüglich Fortführung und Optimierung des Programms ermitteln. Funktionen und bisherige Effekte ermitteln. Analyse des Einflusses auf Zielgruppen und des Fortschritts bei der Erreichung von Programmzielen. Eine Priorisierung von Analyseschwerpunkten innerhalb der breiten Themenstellungen ist nicht zu erkennen. Das federführende Ressort wird als Hauptadressat der Evaluation ausgewiesen. Konkrete Vorstellungen oder Pläne, wie andere Akteure die Evaluationsergebnisse nutzen sollen, und wie dies unterstützt wird, werden nicht dargestellt.	3	
N4 Auswahl und Umfang der Informationen	Das Evaluationsdesign basiert, ausgehend von einer breiten Aufgabenbeschreibung, auf zwei qualitativen Methoden. Diese erscheinen grundsätzlich gut geeignet, ermöglichen aber doch nur eine ausschnittshafte Untersuchung. Eine Logic Chart gliedert Ziele, Instrumente, Outputs und Outcomes des Programms. Mit dieser Logic Chart wird, wie die EvaluatorInnen angeben, die Programmlogik von ihnen rekonstruiert. Im Verhältnis zu weiteren textlichen Darstellungen zum Programm, die insgesamt sparsam ausfallen, ergibt sich allerdings, dass die Programmlogik in der Logic Chart offenbar nicht ganz vollständig greifbar wird. In der Aufarbeitung der Programmstruktur wird eine Gruppe von als erwünscht bezeichneten Programmteilnehmern als indirekte Zielgruppe eingeordnet. Es könnte die Frage gestellt werden, ob der systematische Status dieser Gruppe in der Programmlogik damit in jeder Hinsicht ausreichend bestimmt wurde. Diese Gruppe wird in der gesamten Vorgehensweise der Evaluation nicht einbezogen und nicht beleuchtet. Es kann davon ausgegangen werden, dass dadurch Aspekte der Wirklogik des untersuchten Programms bzw. Einflussfaktoren auf die Erreichung von Outcomes und Impacts nicht vollständig erkannt werden konnten. Es finden sich auch Passagen, die darauf hindeuten, dass sich das Interesse der Untersuchung unter anderem darauf richtete, ob die dem Programm zugrunde gelegte Zielgruppe mit den angenommenen Bedarfslagen und Voraussetzungen in dieser Form existiert. Dies würde zu einer guten Einschätzbarkeit auf eine Bedarfserhebung bzw. eine flächendeckend-systematischen Untersuchung von Strukturvoraussetzungen verweisen. Einige relevante Hinweise werden aus Angaben der untersuchten Zielgruppen gewonnen. Umfassend-systematische Untersuchungen anhand objektiver Daten wurden jedoch nicht angestellt. Auf die Heranziehung verfügbarer Informationen über das untersuchte Programm wurde teilweise verzichtet. Zwischen-evaluationen von innerhalb des Programms geförderten Projekten	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	wären vorgelegen, wurden jedoch nicht herangezogen. Die Evaluation könnte sich hier der Möglichkeit begeben haben, bereits sichtbare Effekte zu erkennen und relevante Einsichten in die konkreten Wirkungsweisen des Programms samt etwaigen Einflussfaktoren zu gewinnen. Sollte dies aus der Art der Zwischenevaluationen heraus nicht möglich gewesen sein, so wäre es angebracht gewesen, dies als Begründung einer methodischen Entscheidung zu berichten (vgl. G3). Es wird dargestellt, dass die Konzeptualisierung des untersuchten Programms auf eine Geschichte von Vorgängerprogrammen zurückgeht und in seiner Wirkungsweise im Verhältnis zu anderen Programmen konzeptualisiert ist. Aussagekräftige Unterlagen über diese anderen Programme (z.B. frühere Evaluationen), aus denen sich unter Umständen weitere Hinweise für die spezifische Wirkungsweise des untersuchten Programms ergeben hätten können, wurden nicht systematisch aufgearbeitet.	
N5 Transparenz von Werten	Der dargestellte Evaluationsauftrag beinhaltet die Vornahme von Bewertungen, ohne dass ersichtlich würde, ob bzw. dass dabei auch Bewertungsmaßstäbe benannt worden wären. Die Evaluation stützt sich auf eine Auswertung von Programmdokumenten sowie eine Abbildung der Sichtweisen von unterschiedlichen Akteursgruppen. Indem eine von drei Gruppen von potenziellen Programmteilnehmern (Zielgruppen) aus der angestellten Untersuchung ausgespart wurde, entsteht dennoch kein völlig ausgewogenes Bild. Für die Einschätzbarkeit des Programms wird eine Reihe von Indikatoren benannt, die offenbar aus dem programminternen Monitoring stammen oder von diesem angewendet werden könnten. Diese Indikatoren werden jedoch nicht eingesetzt. Hier ist zuzugestehen, dass dies zum Evaluationszeitpunkt auch nur sehr eingeschränkt möglich war. Dieser Umstand verbleibt jedoch implizit und muss von den BerichtsadressatInnen selbst geschlossen werden. Eine Unschärfe von Programmformulierungen wird festgestellt und die EvaluatorInnen nehmen einige Begriffsklärungen vor, um Zielformulierungen handhabbar zu machen. Die eingesetzten Begrifflichkeiten nehmen in der Folge den Status von Bewertungsmaßstäben für das Programmkonzept und die Mitteleinsätze in Bezug auf intendierte Ziele an. Es wird nicht transparent gemacht, auf wessen Sichtweisen die gewählten Definitionen zurückgehen. Es wird auch nicht darauf eingegangen, dass bzw. inwiefern sie Auswirkungen auf die erbrachten Ergebnisse gehabt haben können.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist klar und übersichtlich. Er enthält einige wenige Grafiken, die in Bezug auf die vermittelten Berichtsinhalte auch ausreichend für eine gute Inhaltsvermittlung erscheinen. Ein Executive Summary ist nicht beigegeben. Ergebnisse von Untersuchungsschritten werden in einem zusammenfassenden Stil dargestellt. Die Faktenlage wird nicht im Detail präsentiert und wird nicht vollständig transparent. In manchen Berichtsabschnitten fehlen Angaben, durch welche Untersuchungsschritte die berichteten Fakten gewonnen wurden. Der Bericht hat den Charakter eines Kurzberichts und es ist nicht entnehmbar, ob auch eine längere Version erstellt wurde. Es kann nicht beurteilt werden, welche Einschätzung die Berichterstattung bzw. Evaluation, über die berichtet wird, anhand eines solchen Langberichts verdient hätte.	3
N8 Nutzung und Nutzen der Evaluation	Aus Angaben zur Einbeziehung von Beteiligten und Betroffenen (N1) kann geschlossen werden, dass mehrere relevante FTI-politische Akteure auch jenseits der unmittelbaren Programmzuständigkeit sowie Zielgruppen auf die Evaluation aufmerksam wurden, sodass sie in der Folge Interesse an ihren Ergebnissen haben konnten. Ob und inwiefern die FTI-politischen Akteure auch in die Konzeption der Evaluation eingebunden waren, sodass sie ihren jeweiligen Informationsbedürfnissen entgegen kommt, wird nicht dargestellt. Hinweise auf konkrete Vorstellungen über genauere Nutzungsweisen oder Schritte zu ihrer Förderung und Unterstützung finden sich nicht. Inwiefern die Programmverantwortlichen und FTI-politische	3

	Akteure, von denen angenommen werden kann, dass sie über das Programm bereits gut informiert sind, durch die konkreten Berichtsinhalte tatsächlich Lernangebote und Entscheidungshilfen erhalten haben, muss dahin gestellt bleiben. Einschränkungen für die Nützlichkeit des Berichts ergeben sich aus dem Erfüllungsgrad mehrerer Nützlichkeits-, Fairness und Genauigkeitsstandards.	
F3 Vollständige und faire Überprüfung	Der Bericht ist in einer neutralen Sprache abgefasst. Die Sinnhaftigkeit der Programmanlage wird anhand von Sichtweisen von Fördernehmern bzw. Mitgliedern von Zielgruppen analysiert. Die EvaluatorenInnen übernehmen diese Sichtweisen direkt, ohne dazu Stellung zu nehmen, ob und inwiefern sie diese Sichtweisen auch anhand anderer Informationen, die ihnen zur Verfügung stehen bzw. stehen hätten können (vgl. N4), in jeder Hinsicht teilen. Die Sichtweisen einer der drei Zielgruppen werden dabei nicht erhoben und folglich auch nicht zur Geltung gebracht. Der Bericht konzentriert sich streckenweise auf Merkmale der Programmanlage, die von den EvaluatorenInnen als konstruktiv wahrgenommen werden. Er ist in diesem Sinn ein Stärken-Bericht, der hier etwaige geortete Schwächen der bisherigen Programmanlage nicht explizit macht. In anderen Berichtsabschnitten stehen identifizierte Schwächen im Vordergrund, die zum Anlass von Empfehlungen zur Adjustierung des Programms werden, aber doch auch im Verhältnis zu Weichenstellungen bei herangezogenen Informationen (N4) und Bewertungsmaßstäben (N5) zu begreifen sind.	4
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor.	3
G2 Kontextanalyse	In der Programmbeschreibung wird festgehalten, dass das untersuchte Programm durch andere Programme kontextualisiert ist, die auf ähnliche Bedarfslagen zielen und überlappen können. Hierzu finden sich grobe Skizzen, die aus dem Hintergrundwissen der EvaluatorenInnen gespeist sind. Die förderpolitische Positionierung wird jedoch nicht gezielt anhand von Daten untersucht. Hier ist allerdings zuzugestehen, dass die faktische Positionierung für das untersuchte Programm zum Evaluationszeitpunkt nur sehr eingeschränkt untersuchbar war. Es wird weiterhin festgehalten, dass das Programm auf regionaler Ebene wirken soll und auch auf regionale Voraussetzungen angewiesen ist, um die angestrebten Wirkungen zu erreichen. Dies wird nicht untersucht. Es wird nicht thematisiert, welche Faktoren im Bereich der drei Zielgruppen außerhalb der unmittelbaren Reichweite des Programms dazu beitragen können, die beabsichtigten Wirkungen zu fördern oder zu behindern, sodass es angeraten erscheinen könnte, diese Faktoren bei der Konzeptualisierung des Programms und bei der Gestaltung seines Mitteleinsatzes stärker zu berücksichtigen. Die skizzenhafte Abhandlung der Positionierung im Förderportfolio vernachlässigt die Tatsache, dass zwei der drei Zielgruppen auch Gegenstand institutioneller Finanzierung und Steuerung sind. Dadurch kann etwaiges Potenzial, die angestrebten Effekte auch an anderen Einsatzpunkten des FTI-politischen Portfolios mit zu befördern oder gleichzeitig zu behindern, unerkannt geblieben sein.	4
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht stellt die eingesetzten Methoden einleitend dar und gibt in Anhängen ausreichende Informationen, um Vor- und Nachteile dieser Vorgehensweise der Evaluation einschätzen zu lassen. Das untersuchte Programm wird soweit dargestellt, dass auch bislang nicht Vertraute der Berichterstattung über die Evaluationsergebnisse folgen können. Eine wesentliche Rolle für die Vorgehensweise der Evaluation nehmen die Rekonstruktion der Programmlogik in der Logic Chart	4

	und eine Reihe von durch die EvaluatorInnen vorgenommenen Definitionen ein, mit denen Programmformulierungen interpretiert werden. Die Tatsache, dass die EvaluatorInnen für die Untersuchung leitende Definitionen vornehmen, wird explizit gemacht. Eine Abhaltung eines Logic Chart-Workshops mit Programmbeteiligten wird nicht erwähnt. Es wird nicht vermittelt, inwiefern es sich in beiden Fällen um Sichtweisen der EvaluatorInnen handelt, die sie von außen an das Programm anlegen, oder um von Programmbeteiligten mitgetragene Sichtweisen. Insgesamt werden wesentliche Entscheidungen in der Vorgehensweise nicht gut transparent gemacht, um alle BerichtsadressatInnen jedenfalls erkennen zu lassen, wie diese Schritte gesetzt wurden, und welche Auswirkungen sie auf die Ergebnisse und Schlüsse gehabt haben können. Verfügbare Informationen über das bisher ausgelöste Projektgeschehen wurden nicht herangezogen, ohne dafür eine Begründung zu geben. Es wird nicht dargestellt, welche der eingangs genannten Themenstellungen für die Evaluation durch die eingesetzten Methoden wie gut beleuchtet werden konnten. Die auf spärlichem Methodeneinsatz basierte Evaluation hätte allerdings speziellen Anlass gegeben, diesen Punkt in allgemein verständlicher Weise zu beleuchten.	
G8 Begründete Schlussfolgerungen	Schlussfolgerungen und Empfehlungen werden in einem eigenen Kapitel getrennt von Ergebnisdarstellungen präsentiert. Schlussfolgerungen und Empfehlungen sind dabei formal voneinander getrennt, allerdings führen auch die Schlussfolgerungen Empfehlungen mit. Es wird eine Reihe von spezifischen, umsetzbaren Empfehlungen zur Programmumsetzung ausgesprochen. Darüber hinaus werden Empfehlungen zur grundsätzlichen Programmgestaltung abgegeben. Hier wird allerdings teilweise mit angenommenen Wirklogiken und Effekten argumentiert, die in dieser Form zuvor nicht berichtet bzw. nicht untersucht wurden. Dies kann unter Umständen darauf zurück zu führen sein, dass neben dem vorliegenden Kurzbericht auch ein Langbericht ausgearbeitet wurde, der die entsprechenden Angaben enthält. Das Vorliegen eines Langberichts ist jedoch nicht gesichert. Hinweise auf die Herkunft bestimmter Sichtweisen auf das Programm, die sich im vorangegangenen Bericht finden, werden im Schlusskapitel nicht mehr mitgeführt. Es werden Weichenstellungen für das Programm empfohlen, die substanzielle Eingriffe in die bisherige Programmformulierung darstellen würden. Diese Empfehlungen werden nicht mit Hinweisen auf alternative Optionen versehen, die bei einer Klärung von diagnostizierten Unschärfen der bisherigen Programmformulierung zur Verfügung stehen könnten. Es werden auch nicht potenzielle Nachteile herausgearbeitet, die sich aus den vorgeschlagenen Adjustierungen für die Gesamt-Wirklogik des Programms ergeben könnten, je nach deren Auffassungsweise bzw. diesbezüglichen Entscheidungen.	4

Evaluationsbericht Nr. 10		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	summativ (ex post-Evaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Programmverantwortliche, Programmumsetzende, und vom Programm Betroffene (Programmtteilnehmer und weitere Teile der Zielgruppen) wurden im Rahmen von Datenerhebungsverfahren einbezogen. Des Weiteren wurden in Datenerhebungsverfahren einige Akteure der FTI-Politik einbezogen, die in die Konzeption des untersuchten Programms eingebunden oder im Umfeld des Programms tätig waren. Es kann allerdings nicht entnommen werden, ob abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene in die Planung der Evaluation hinsichtlich von Informationsbedürfnissen einbezogen wurden.	2	
N2 Klärung der Evaluationszwecke	Angegeben wird: Abschließende Bewertung des Programms und Erbringen von Hinweisen für die Gestaltung rezenter und zukünftiger Fördermaßnahmen. Unter Einsatz eines geeigneten Methodenportfolios die Einbettung des Programms in das österreichische Innovationssystem sowie Effekte und Wirkungen bewerten. Die bisherigen Erfahrungen zusammenfassen und kritisch neu bewerten, bislang unberücksichtigte Bereiche stärker beleuchten und Schlüsse ziehen, die für eine Einschätzung des Programmgegenstands in Österreich notwendig und hilfreich sind. Es wird sodann eine Hauptfrage in den Mittelpunkt der Untersuchung gestellt. Damit liegt eine gewisse Priorisierung innerhalb der sonst recht global beschriebenen Zwecksetzung vor. Allerdings ist auch diese Fragestellung von umfassender Art (vgl. G2).	2	
N4 Auswahl und Umfang der Informationen	Evaluationsfragestellungen werden benannt. Zu deren Beantwortung wird ein Mix aus quantitativen und qualitativen Methoden eingesetzt, der als gut geeignet erachtet werden kann. Die gewählte Analysestrategie besitzt eine Reihe von Stärken: • Die Interventionslogik des untersuchten Programms wird aufgeschlüsselt und in direkten Konnex mit detaillierteren Evaluationsfragen gebracht. In der Aufschlüsselung der Interventionslogik benennt der Bericht klar die ursprüngliche Programmanlage als Bezugspunkt. • Die Evaluation setzt unter anderem eine avancierte Methode ein, die z.B. auch von der Weltbank für entsprechende Zwecke empfohlen wird. Mit dieser Methode werden Schwierigkeiten bei der Etablierung eines echten Kontrollgruppendesigns in einer bestimmten Weise bewältigt. • Fachliteratur und anderweitig erstellte Berichte zum Evaluationsgegenstand bzw. zum Operationsbereich des Programms werden in substanzieller Weise herangezogen. • Die Untersuchung verfährt in Abschnitten theoriegeleitet. • In der konzeptiven Auseinandersetzung mit Grundannahmen des Programms wird es möglich, Wirkmechanismen zu benennen, die für die Erreichung des Programmerfolgs als entscheidend erachtet werden können, und bestimmte Handlungsweisen gegenüber diesen Mechanismen als wesentliche Bedingung für die Erreichung der Programmziele auszuweisen. • Verfügbare offizielle Datenquellen werden umfänglich genutzt. • Ergebnisse mehrere Methoden werden gespiegelt, um eine Beleuchtung derselben Phänomene von hoher Tragfähigkeit zu erzielen. • Einzelne Schritte der vorgenommenen Untersuchung gehen auch	2	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	möglichen nicht-intendierten Effekten des Programms nach. Es sind allerdings auch Schwächen zu verorten: - Die geplante Analyse auf Basis einer der eingesetzten Methoden, die tragende Rolle für Teile der Gesamtanalyse hat, wird nur soweit vorangetrieben, dass dies eher als Vorbereitungsschritt für eine daran anschließende eingehendere Analyse erscheint, die dann die dargestellte Absicht tatsächlich erfüllt hätte. Es wird eine Typologisierung angestrebt, die sich allerdings in der Aneinanderreihung von Einzeldimensionen erschöpft, die als solche kaum aussagekräftig sind. Die Analyse hätte weiter vorangetrieben werden sollen, oder die kaum aussagekräftigen Ergebnisse einer unvollständig durchgeführten Analyse hätten nicht in der vorliegenden Form berichtet werden sollen (vgl. N6, G3). Diese Beobachtung kann einen Hinweis darstellen, dass der Evaluation nicht ausreichend Ressourcen für die Durchführung eines grundsätzlich iterativen und anspruchsvollen Verfahrens zur Verfügung gestanden haben. Alternative methodische Vorgehensweisen, die nicht in einem anspruchsvollen Rechenverfahren bestanden hätten, wurden nicht gewählt. - Die Evaluation stützt sich auf erhältliche Monitoringdaten, wobei die Programmschienen unterschiedlich gut dokumentiert und somit Daten lückenhaft sind, die für die Evaluation von zentralem Interesse gewesen wären. Durch zusätzliche Eigenerhebungen hätte sich die Evaluation hier eine noch bessere, gleichmäßige Datenbasis verschaffen können.	
N5 Transparenz von Werten	Der dargestellte Evaluationsauftrag beinhaltet die Vornahme von Bewertungen, ohne dass ersichtlich würde, ob bzw. dass dabei auch Bewertungsmaßstäbe benannt worden wären. In der konkreten Analysestrategie wählen die EvaluatorenInnen den Bezugspunkt des Systemversagens im Nationalen Innovationssystem. Dadurch ergibt sich ein konzeptiver Konnex unter allen verfolgten Fragestellungen. Um das Programm einschätzen zu können, werden objektive Daten ebenso wie Sichtweisen von Programmteilnehmern und Zielgruppen herangezogen. Einflüsse auf die Einschätzung des Programms können sich aus der Wahl bestimmter Beschreibungsweisen und Indikatoren ergeben haben, die letztlich auf ihre epistemologische Grundlage zu untersuchen wären, um ihre genaue „Brille der Wahrnehmung“ bestimmen zu können. Dies kann nicht als Aufgabe einer Evaluation verstanden werden, und die eingesetzten Begrifflichkeiten und Messgrößen bewegen sich im Rahmen des allgemein Anerkannten. Allerdings enthält der Bericht doch auch keine Hinweise, ob und inwiefern auch andere Begrifflichkeiten und Messgrößen im Rahmen des verfolgten Ansatzes bestehen hätten können, bzw. ob sie zu einer anderen Einschätzung führen hätten können. (Ein Beispiel wäre hier etwa der Unterschied zwischen quantitativ gefasster Kooperationsintensität und qualitativen Kooperationsmerkmalen.)	2
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht hat einen strikt logisch-hierarchischen Aufbau, dem leicht zu folgen ist. Präsentationen und Diskussionen sind ausführlich und schrittweise gestaltet, was sehr gute Nachvollziehbarkeit gewährleistet, wenn auch damit den AdressatInnen einiges an Lesearbeit abverlangt wird. Die Abfassung eines Kurzberichts, von dem aus in die ausführlichen Darstellungen hineingegriffen werden kann, hätte die Handhabbarkeit des Berichts für „gestresste“ AdressatInnen erhöhen können. Grafiken und Tabellen veranschaulichen konzeptive Inhalte sowie Ergebnisse. Ein Grafik- und Tabellenverzeichnis, anhand dessen sie rasch auffindbar werden, wurde allerdings nicht beigegeben. Einzelne Grafiken und Tabellen hätten im Interesse der raschen Auffassbarkeit noch besser beschriftet werden können. Ein Executive Summary ist beigegeben und auch in englischer Sprache vorhanden. Es erfolgt eine klare und eingehende Darstellung des Programms in seinem Entstehungskontext, die auch bislang nicht mit dem Programm Vertrauten ermöglicht, den Evaluationsgegenstand zu	2

	verstehen und der Berichterstattung über die Evaluation zu folgen. Verschiedene Berichtsabschnitte sind auf unterschiedlichem sprachlichen Niveau abgefasst. In einem Abschnitt wird eine theoretische Analyse vorgelegt, die in anspruchsvoller Weise in wissenschaftlichem bzw. auch lehrbuchartigem Duktus präsentiert wird. Diese Präsentationsweise könnte Gruppen von BerichtsadressatInnen, die nicht laufend mit wissenschaftlichen Diskussionen umgehen, überfordern. Um die Auffassbarkeit zu erhöhen, hätten Teile in Fußnoten oder einen Anhang ausgelagert werden können. In einem Berichtsabschnitt werden die LeserInnen in eine umfängliche Präsentation von wenig aussagekräftigen Daten hineingezogen. Hier wäre eine deutliche Raffung und Konzentration auf den wesentlichsten Gehalt besser gewesen.	
N8 Nutzung und Nutzen der Evaluation	Zentrale Programmbeteiligte und einige Akteure der FTI-Politik die in die Konzeption des untersuchten Programms eingebunden und/oder im Umfeld des Programms tätig waren, wurden in die Evaluation einbezogen (vgl. N1). Daneben wurden Zielgruppen des Programms im Rahmen von Datenerhebungen auf die Evaluation aufmerksam gemacht, sodass sie ein Interesse an deren Ergebnissen entwickeln konnten. Anhand der vorliegenden Angaben kann nicht festgestellt werden, welche Bereiche der FTI-Politik auf welchen Verantwortungsebenen einbezogen waren. Hinweise auf konkrete Vorstellungen über genauere Nutzungsweisen oder Schritte zu ihrer Förderung und Unterstützung finden sich nicht. Es liegt ein Bericht vor, bei dem jedenfalls davon ausgegangen werden kann, dass er seinen AdressatInnen zu neuen Einsichten verhelfen und sie bei zukünftigen Konzeptualisierungen unterstützen kann. Leichte Einschränkungen für die Nützlichkeit des Berichts ergeben sich aus mehreren Nützlichkeits-, Fairness- und Genauigkeitsstandards, die jeweils theoretisch noch besser erfüllt werden hätten können.	2
F3 Vollständige und faire Überprüfung	Der Bericht ist einer neutralen Sprache abgefasst, wie sie für wissenschaftliche Arbeiten typisch ist. Er benennt sowohl Stärken als auch Schwächen des untersuchten Programms. Sichtweisen von Fördernehmern, Zielgruppen und Programmverantwortlichen gehen in die Analyse ein und werden so präsentiert, dass die Herkunft der jeweiligen Angaben gut nachvollziehbar ist. Die textliche Umsetzung bzw. Interpretation von quantitativen Ergebnissen ist an der Vollständigkeit des jeweiligen Gehalts orientiert. Qualitative Ergebnisse werden anhand ausgewiesener Zitate gut nachvollziehbar und werden auch im Sinne von Gesamtverteilungen greifbar gemacht. Abstriche ergeben sich allerdings aus einem Berichtsabschnitt, der sich deutlich neben beobachteten Fakten auch auf Annahmen stützt, deren Herkunft nicht transparent gemacht wird.	2
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das Programm wird in seiner Stellung zum Nationalen Innovationssystem begriffen und erläutert, und es werden in dieser Hinsicht relevante Fragestellungen untersucht. Die durchgeführte Analyse hat allerdings ihren Schwerpunkt in der Aufarbeitung von Effekten, die unter unmittelbar Begünstigten entstanden sind. Ein über die unmittelbaren Akteure hinausreichender Programmnutzen wird von den EvaluatorsInnen zwar als prinzipiell wesentliche Komponente eines Programmserfolgs ausgewiesen, konnte aber anhand der eingesetzten Methoden nicht erfasst werden. Es hätten noch weitere Erhebungen und Analysen angestellt werden können, um über das dem Programm unmittelbar	3

	zurechenbare Geschehen hinaus auch eventuelle Hinderungsfaktoren für das Eintreten längerfristig erwarteter Gesamtwirkungen für das Innovationssystem erkennen zu können, wobei auch Folge- und Nebenwirkungen des untersuchten Programms relevant sein können. Es kann davon ausgegangen werden, dass die ausgewiesene Hauptfragestellung, ob das geschaffene Anreizsystem auch aktuell die adäquate Antwort auf die Strukturherausforderungen des Innovationssystems darstellt, noch weitere Fragen in sich birgt, die den gezielten Umgang mit hoher Komplexität erfordern. In der vorliegenden Analyse wird greifbar gemacht, welche Merkmale das vom Programm geförderte Geschehen hat, jedoch nur recht bedingt, welche Beiträge dem Programm zur übergreifenden Erreichung der ihm zugrunde liegenden Zielsetzungen insgesamt im Innovations-system zugeschrieben werden können.	
G3 Beschreibung von Zwecken und Vorgehen	Ausgehend von der globalen Zwecksetzung (vgl. N2) benennt der Bericht einige Evaluationsfragestellungen, die die Untersuchung strukturieren. In einem kurzen Methodenkapitel wird das Methodenset überblicksartig beschrieben und das Rationale des Methodeneinsatzes benannt. Methodische Angaben erfolgen sodann in den einzelnen Berichtsabschnitten. Hier werden zu den meisten Methoden differenzierte Erläuterungen beigebracht, die den Berichts-adressatInnen erlauben, das Wie und Warum der Vorgehensweise zu den verfolgten Fragestellungen verstehen. Unter anderem werden Abwägungen zur Verwendung alternativer Datenbasen dargestellt. Zu manchen Analyseschritten, jedoch nicht allen, werden Einschränkungen für die Analysierbarkeit und für die Aussagekraft von Ergebnissen offengelegt. Die methodische Diskussion setzt sich mit der Kausalitätsproblematik auseinander und weist auf Einschränkungen der Interpretationsmöglichkeiten der verfügbaren Daten hin. In einem Berichtsabschnitt werden jedoch keine ausreichenden Angaben gemacht, um das Vorgehen für BerichtsadressatInnen transparent zu machen. Es unterbleibt auch ein jedenfalls allgemein verständlicher Hinweis darauf, dass streckenweise Daten analysiert werden, die nur bedingt die Anforderung für eine aussagekräftige Wirkungsanalyse erfüllen, einen sachlich benötigten Zeitabstand zu auslösenden Programmereignissen aufzuweisen. Hier ist nicht ausgeschlossen, dass auch verfehlte Auffassungen bei BerichtsattInnen ausgelöst werden könnten. Über weite Strecken sind Daten und Interpretationen so angeboten, dass wissenschaftlich gebildete LeserInnen Stärken und Schwächen der Methodik und der durch sie erbrachten Ergebnisse erkennen können. Es kann allerdings nicht davon ausgegangen werden, dass die vorliegenden Darstellungen es auch anderen Adressatengruppen leicht machen, etwaige Vor- und Nachteile des Evaluationsdesigns zu verstehen. Eine allgemein verständliche Überblicksdarstellung, in der Vor- und etwaige Nachteile des Designs erläutert werden, findet sich nicht.	2
G8 Begründete Schlussfolgerungen	Schlussfolgerungen finden sich am Ende von Einzelkapiteln und sind von Faktendarstellungen durchwegs klar getrennt. Diese Schlussfolgerungen sind im Bezug auf die zuvor dargelegten Informationen und Interpretationen im Allgemeinen gut nachvollziehbar. In einem Resümee scheint eine Diskrepanz zwischen einem hier genannten Datum und dem zuvor Berichteten aufzutreten. Die Funktion von Gesamtschlussfolgerungen wird vom Executive Summary übernommen, das zunächst die zentralen Berichtsinhalte in holistischer Weise sowie unter Benennung zentraler Fakten zusammenfasst. Im Anschluss werden einige Empfehlungen ausgesprochen, die allerdings nicht als solche gekennzeichnet sind. Diese sind teils in einem Nachfolgeprogramm umsetzbar, indem eine Bewältigung prinzipieller Herausforderungen durch adjustierte Vorgehensweisen in der Programmumsetzung das Leitmotiv darstellt. Im Umgang mit groß dimensionierten Fragen, welche förder-	2

	politischen Aufgaben und Kautelen sich zum aktuellen Zustand des Innovationssystems ergeben können, geht die Berichterstattung hier anhand prinzipiell-theoretischer Überlegungen auch über klar belegte Wirkmechanismen hinaus.	
--	--	--

Evaluationsbericht Nr. 11		Beobachtungszeitraum: 2007 - 2010	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Effektivität		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen der Datenerhebung wurden Fördernehmer (Organisationen und deren VertreterInnen) einbezogen. Dabei wurden in erster Linie direkte Zielgruppen (Antragsberechtigte bzw. an der Umsetzung geförderter Projekte Beteiligte) berücksichtigt. Eine mögliche Einbeziehung der indirekten Zielgruppe, die als die eigentliche Gruppe der zentralen Begünstigten des Programms zu erachten ist, wird ebenfalls erwähnt, ist aber nicht konkret nachvollziehbar. Es wurde ein Verfahren angewendet, in dem zunächst angesprochene Betroffene andere Betroffene kontaktierten, um eine breite und gute Abbildung von unterschiedlichen für Programmgeschehen und Programmerfolg relevanten Gruppen zu ermöglichen. Es ist nicht erkennbar, ob und gegebenenfalls wie Programmverantwortliche und -umsetzende über das bloße Faktum der Auftragserteilung hinaus einbezogen wurden, um Informationsbedürfnisse zu klären. Es findet sich auch kein gesicherter Hinweis, dass andere FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben könnten.	3	
N2 Klärung der Evaluationszwecke	Angegeben wird: Evaluation gemäß den dem Programm zugrunde liegenden Sonderrichtlinien. Erfolgskontrolle des Programms anhand vorab festgelegter Indikatoren, um auf dieser Grundlage die inhaltliche Ausrichtung, die Abwicklung und die Bewertungskriterien des Programms in seiner nächsten Programmphase zu gestalten. Es werden zwei unterschiedliche Evaluationszwecke, Erfolgskontrolle und Ausrichtung auf instrumentellen Nutzen, miteinander verbunden. Eine Priorisierung zwischen diesen in der Aufgabenbeschreibung der Evaluation ist nicht zu erkennen. In der Angabe bleibt auch unklar, ob mit „Bewertungskriterien des Programms“ Kriterien gemeint sind, die innerhalb des Programms bei der Projektauswahl zur Anwendung gelangen, oder Bewertungskriterien, die an das Programm hinsichtlich seines Gesamterfolgs angelegt werden sollen.	3	
N4 Auswahl und Umfang der Informationen	Der Evaluationszeitpunkt ist früh gewählt, sodass das Programm nur ausschnitthaft, mit Schwerpunkt auf einer seiner beiden Förderschiene, untersucht werden kann. Genauere Evaluationsfragestellungen, anhand derer die Analyse strukturiert wird, werden nicht benannt. Ausgehend von den global formulierten Zwecksetzungen bringt die Evaluation zwei Methoden zur Anwendung, die grundsätzlich als gut geeignet gelten können. Eine quantitative Analyse, die der vorab festgelegten Betrachtung anhand von einigen Indikatoren entspricht, wird durch den Einsatz einer qualitativen Methode ergänzt, die zusätzliche Möglichkeiten eröffnet, um wertvolle Erkenntnisse über die Programmanlage und -abwicklung zu erbringen, aus denen für die künftige Ausgestaltung gelernt werden kann. Das Programm ist so gestaltet, dass es eine Vielzahl von Einzelprojekten auslösen und über diese unterschiedlichen Einzelprojekte in Richtung seiner selbstgesetzten Ziele wirken will. Die grundsätzliche Entscheidung, auf das Projektgeschehen differenziert einzugehen, ist als sachdienlich zu erachten. Dass hier lediglich ein Teil des Projektgeschehens analysiert wird, kann als sachgerecht erachtet werden. Es wird jedoch keine Strategie angegeben, wie der Ausschnitt bestimmt wird, und es findet sich ein Hinweis, dass hier auch Bequemlichkeitsgründe mitgespielt haben könnten.	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Die Programmlogik wird nicht strukturiert aufgearbeitet. Nicht an allen Stellen der Dateninterpretation wird klar, ob in der Auseinandersetzung mit Projektformen und mit Mechanismen, die deren Zustandekommen und die daraus ausfließenden Effekte beeinflussen können, Programmoutputs, Outcomes, Effekte im Sinne längerfristiger Zielsetzungen, oder positive Nebenwirkungen verhandelt werden. Verschiedene Hinweise zu Programmmerkmalen an unterschiedlichen Berichtsstellen scheinen darauf hinzudeuten, dass die Wirklogik des Programms nicht hinsichtlich aller Umsetzungsaspekte untersucht wurde. Die Erstellung einer Logic Chart oder eine andere Form der strukturierten Aufarbeitung der Programmstruktur und der logisch-hierarchischen Wirklogik hätte die Analyse jedenfalls unterstützen können. Eine Aufgabenstellung, auch Bewertungskriterien zu reflektieren, die innerhalb des untersuchten Programms bei der Projektvergabe angewendet werden (vgl. N2), hätte eine eingehende und gut strukturierte Analyse der intendierten und faktischen Wirklogik des Programms angelegen erscheinen lassen, da hier die logischen Voraussetzungen für die Wahl von Selektionskriterien für vielversprechende Projektformen anzunehmen sind. Dies findet sich jedoch nicht. Die EvaluatorInnen präsentieren einige Daten aus einer statistische Auswertung von Monitoringdaten, die gemäß ihren Darlegungen zum Evaluationszeitpunkt nur eingeschränkt analysiert werden können. Einer Aufgabenstellung, Bewertungskriterien für das Gesamtprogramm zu reflektieren (vgl. N2), wird mit einer Reflexion der Aussagekraft einzelner Kennzahlen teilweise entsprochen. Eine vollständige Auswertung zu allen Indikatoren wird dabei nicht vorgelegt. Es wird auch nicht herausgearbeitet, welches Set an Kennzahlen von den EvaluatorInnen als künftig sinnvoll anwendbar erachtet wird. Zwei von unterschiedlichen Methoden getragene Berichtsteile stehen isoliert nebeneinander. Eine Beziehung zwischen quantitativen und qualitativen Ergebnissen wird nicht hergestellt. Die angewendete Vorgehensweise ist nur eingeschränkt geeignet, um auch Faktoren erkennen zu können, die Teile der für die Auslösung des intendierten Projektgeschehens wesentlichen Zielgruppen bisher davon abgehalten haben, am Programm teilzunehmen, sodass Möglichkeiten erkannt werden könnten, wie das Programm diesen Gruppen durch Adjustierungen künftig besser begegnen könnte. Erhebungen unter bisherigen Nicht-Teilnehmern hätten hier zu einer besseren Informationsbasis führen können. Die Bedeutung von Merkmalen des Projektgeschehens für die indirekte Zielgruppe, in der die erwarteten Wirkungen des Programms eintreten sollen, hätte durch gezielte Erhebungen in dieser Gruppe deutlich besser beleuchtet werden können. Es erfolgt an manchen Stellen eine Bezugnahme auf die Inhalte von einigen geförderten Einzelprojekten, die als sachdienlich zu erachten ist. Eine systematische Analyse der wohl flächendeckend verfügbaren Projektbeschreibungen ist jedoch nicht erfolgt. Das von der Evaluation betrachtete Projektgeschehen hätte auch in Form von Fallstudien (Case Studies) analysiert werden können, was Vorteile hinsichtlich einer systematisch strukturierten und vergleichenden Analyse haben hätte können, in der die ausgelösten Projekte entlang wesentlicher Merkmale wirklich greifbar werden.	
N5 Transparenz von Werten	Es finden sich keine Angaben über Bewertungsmaßstäbe, die zur Einschätzung des Programms herangezogen werden. Mangels einer klaren Darstellung der Programmziele, der vom Programm eingesetzten Mittel zur Erreichung dieser Ziele, oder von Evaluationsfragen kann von BerichtsadressatInnen kaum eingeschätzt werden, ob die Einschätzung des Programms in adäquater Weise erfolgt. Die Indikatoren, die in der Angabe des Evaluationszwecks firmieren, werden nicht tragend für die vorfindliche Analyse. Förderliche und hinderliche Faktoren für Wirkungsweisen des Programms werden im qualitativen Teil der Untersuchung auf der Basis von Erfahrungen und Sichtweisen von Gruppen identifiziert, denen in der Programmumsetzung eine wesentliche Rolle zukommt. Hierbei wird allerdings nur an einer Berichtsstelle zwischen Sichtweisen unterschiedlicher Herkunft differenziert und	4

	im Übrigen keine Auseinandersetzung dazu angeboten, inwiefern die erhobenen Sichtweisen kongruent oder divergent sein könnten, und welche Bedeutung dies für die Umsetzung und Zielerreichung des Programms haben könnte. Die Sichtweise der zentralen Begünstigten des Programms wird nur in unklarer und kaum nachvollziehbarer Weise verfolgt. Es wird angegeben, dass auch diese Gruppe prinzipiell die Möglichkeit hatte, an den Erhebungen teilzunehmen. Es wird jedoch nicht gut erkennbar, inwieweit Aussagen aus dieser indirekten Zielgruppe in die Analyse eingehen, und inwiefern Aussagen über die zentralen Begünstigten von anderen Seiten herangezogen werden.	
N6 Vollständigkeit und Klarheit der Berichterstattung	Es liegt ein kurzer Bericht in allgemein verständlicher Sprache vor, der gemäß Auskunft der Auftraggeberorganisation den Gesamtbericht zu dieser Evaluation darstellt. Ein Executive Summary ist nicht beigegeben. Grafiken oder Tabellen werden nicht beigegeben. Beim Umgang mit Indikatoren bzw. zur Darstellung wesentlicher quantitativer Ergebnisse wäre eine grafische oder tabellarische Darstellungsweise jedenfalls sinnvoll gewesen. Das untersuchte Programm wird nur skizzenhaft vorgestellt. BerichtsadressatInnen, die bislang nicht mit ihm vertraut sind, werden nicht dabei unterstützt, das untersuchte Programm gut zu verstehen und der Berichterstattung der Evaluation gut folgen zu können. Die Indikatoren, die für die Einschätzung des Programms vorgesehen sind, werden nicht vollständig und klar dargestellt. Beim Umgang mit quantitativen Daten wird nicht klar kommuniziert, was aktuelle Werte für ursprünglich vorgesehene Indikatoren sind, und was darüber hinaus festgestellte Verteilungen. Die Berichterstattung über qualitative Daten vernachlässigt klare Differenzierungen, welche Effekte tatsächlich eingetreten sind und welche darüber hinaus vermutbar erscheinen, bzw. aus der Programmformulierung heraus erwartet werden, ohne allerdings zum Evaluationszeitpunkt beobachtet werden zu können. Es liegt ein Bericht vor, bei dessen Abfassung offenbar ausschließlich an die HauptadressatInnen (AuftraggeberInnen) gedacht wurde, die über das nötige Vorwissen verfügen, um alle Evaluationsergebnisse einzuordnen. Doch auch für die HauptadressatInnen werden essentielle Inhalte nur schlecht auffassbar dargeboten.	4
N8 Nutzung und Nutzen der Evaluation	Aus der Einbeziehung von Betroffenen (vgl. N1) kann geschlossen werden, dass Teile der Gruppen, die für die Gestaltung des bereits zustande gekommenen Projektgeschehens wichtig sind, auf die Evaluation aufmerksam gemacht wurden, sodass sie in der Folge ein Interesse an deren Ergebnissen entwickeln konnten. Weitere Kreise in den Zielgruppen wurden offenbar nicht angesprochen. Zur Einbeziehung der HauptadressatInnen der Evaluation werden keine Angaben gemacht. Während eine aussagekräftige Durchführung der indikatoren-gestützten Analyse wegen des frühen Evaluationszeitpunkts unterbleibt, erbringt die qualitative Analyse Hinweise, die für die HauptadressatInnen nützliche Informationen darstellen können. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden soll, werden nicht dargestellt. Die Ergebnisse werden in einer Weise berichtet, die für andere InteressentInnen als die HauptadressatInnen wenig geeignet ist, und auch für die HauptadressatInnen mit Vorbehalt zu betrachten ist. Deutliche Einschränkungen für die Nützlichkeit des Berichts ergeben sich aus niedrigen Erfüllungsniveaus für mehrere Nützlichkeits-, Fairness- und Genauigkeitsstandards.	4
F3 Vollständige und faire Überprüfung	Der Bericht ist in neutraler Sprache abgefasst. Er leidet jedoch unter: • Schlechter Differenzierung zwischen konkreten Beobachtungen und Vermutungen, die entweder von Zielgruppen übernommen oder von den EvaluatorsInnen selbst angestellt werden. • Weitgehender Enthaltung der EvaluatorsInnen, Stärken und	4

	Schwächen des untersuchten Programms zu benennen. Der Bericht breitet in erster Linie Beobachtungen aus, zu denen nur bedingt klar wird, welchen Stellenwert die EvaluatorenInnen ihnen für den Programmerfolg zumessen. • Präsentation von statistischen Ergebnisse aus den Monitoringdaten weitestgehend ohne dass die EvaluatorenInnen angeben, wie diese Ergebnisse ihrer Meinung nach verstanden werden sollten bzw. können. BerichtsadressatInnen werden mit den referierten Fakten bis auf wenige Hinweise „alleine gelassen“. • Unklarheiten, inwieweit berichtete Stärken auf in der angestellten Untersuchung beobachteten Fakten beruhen oder aus der (kaum dargestellten) Programmanlage als solcher abgeleitet werden.	
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das Programm wird durch die Fokussierung auf Projekte, die in unterschiedlicher Form zustande kommen und in unterschiedlichen Akteurskonstellationen ihren Verlauf nehmen, grundsätzlich in Relation zu den Umwelten begriffen, in denen es operiert. Mangels genauerer Angaben über das Programm und über die methodische Vorgehensweise (vgl. G3) kann nicht eingeschätzt werden, wie umfassend oder ausschnitthaft diese Betrachtung ist.	3
G3 Beschreibung von Zwecken und Vorgehen	Die Vorgehensweise der Evaluation in methodischer und organisatorischer Hinsicht wird in Eckpunkten dargestellt, wobei die Motivation für die Heranziehung der ergänzenden qualitativen Methode erläutert wird. Die Darstellung muss jedoch als unzureichend erachtet werden, um die Vorgehensweise der Evaluation in allen Hinsichten so transparent zu machen, wie es für eine Nachvollziehbarkeit ihres Vorgehens notwendig wäre. Durch das Fehlen einer griffigen Gesamtdarstellung des Programms und von ausgewiesenen Evaluationsfragestellungen wird es erschwert, die Vorgehensweise der Evaluation hinsichtlich ihrer Analysestrategie und ihrer Tragfähigkeit einzuschätzen. Es ist davon auszugehen, dass BerichtsadressatInnen mit Schwierigkeiten konfrontiert sind, berichtete Ergebnisse auf Programmintentionen zu beziehen und ihre jeweilige Relevanz zu erkennen. Die vorgesehenen Indikatoren werden nicht klar vorgestellt und teils ausgewertet, teils hinsichtlich von Auswertbarkeit und Aussagekraft kommentiert. Die quantitativen Daten werden nicht systematisch hinsichtlich ihrer Vollständigkeit dargestellt. Zur Durchführung der qualitativen Methode werden Eckdaten angegeben. Es wird jedoch keine Sampling-Strategie angegeben und nicht berichtet, welche Zielgruppen in welchem Ausmaß erreicht wurden. Es wird auch nicht berichtet, was die an die Zielgruppen gerichteten Fragestellungen waren. Ergebnisse werden nur stellenweise in einer an unterschiedlichen Zielgruppen orientierten Systematik berichtet, ohne dass für eine sonst öfters anzutreffende Außerachtlassung der Differenzierung Gründe angegeben werden. Auf Einschränkungen der Analysierbarkeit des Programms zum Evaluationszeitpunkt wird hingewiesen. In der Interpretation qualitativer Ergebnisse werden allerdings stellenweise Aussagen getätigt, die dieser Tatsache nicht vollständig Rechnung tragen. Es wird nicht explizit gemacht, welche Programmaspekte wie vollständig oder unvollständig untersucht werden konnten, und wie tragfähig die berichteten Ergebnisse sind.	4
G8 Begründete Schlussfolgerungen	In einem Abschlusskapitel werden zentrale Schlussfolgerungen und Empfehlungen angeboten. Die Textierung reflektiert den frühen Evaluationszeitpunkt, zu dem nur bedingt Aussagen über das Programm möglich werden. Die getätigten Aussagen erscheinen in der vorangegangenen Analyse abgestützt, der Bezug auf konkrete Faktenlagen wird jedoch nicht transparent gemacht. Es werden eher Optionen benannt, wie über das Programm gedacht	4

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	werden könnte, um in der Folge zu weiteren Entscheidungen zu gelangen, als spezifische Hinweise zu konkreten Adjustierungen zu bieten, die die EvaluatorInnen als angeraten erachten. Zwischen Aspekten der Umsetzung des gegebenen Programms, übergeordneter Programmkonzeption und Programmeobachtung wird nicht systematisch differenziert. Zwischen Alternativen wird nicht in begründender Weise abgewogen. Daneben flicht der Berichtstext an verschiedenen Stellen Schlussfolgerungen und Empfehlungen unmittelbar in Faktendarstellungen ein, die einzelne Gesichtspunkte der Programmgestaltung betreffen, und macht sie dabei nicht als solche kenntlich. Dabei werden zum Teil Informationen angezogen, die nicht klar belegt werden, oder stellenweise auch Überinterpretationen darstellen könnten.	
--	--	--

Evaluationsbericht Nr. 12		Beobachtungszeitraum: 2011- 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	Formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In Datenerhebungsverfahren wurden Zielgruppen des untersuchten Programms (FördernehmerInnen und nicht erfolgreiche AntragstellerInnen) sowie Verantwortliche in den Organisationen, die den institutionellen Kontext des untersuchten Programms darstellen, einbezogen. Ein Workshop mit den AuftraggeberInnen zur Reflexion von Zwischenergebnissen wurde abgehalten. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber aus ihren Positionen im Steuerungssystem heraus an der Evaluation und ihren Ergebnissen Interesse haben konnten.	3	
N2 Klärung der Evaluationszwecke	Angegeben wird: Kritische Würdigung des Programms mit Schwerpunkt auf der Analyse von Effekten und Wirkungen, wobei sowohl direkte Wirkungen auf FördernehmerInnen als auch indirekte Wirkungen im Programmkontext Beachtung finden sollen. Bewertung bisheriger Leistungen und Wirkungen und Ermittlung handlungsleitender Perspektiven und Gestaltungsinformationen. Erbringung eines Beitrags zur Einschätzung der Aktualität der Programme und zur Unterstützung der strategischen Steuerung und Positionierung der Programme. Zu den recht breiten Zwecksetzungen wird eine Priorisierung der Wirkungsdimension erkennbar. Nicht angegeben werden Vorstellungen oder Pläne dazu, wer die vorgesehenen Ergebnisse in welcher genaueren Weise nutzen wird.	2	
N4 Auswahl und Umfang der Informationen	Die Evaluation setzt vier Methoden zur Datenerhebung und -analyse parallel zueinander ein, die als gut geeignet gelten können. Genauere Evaluationsfragestellungen, anhand derer die Analyse ausgehend von den Evaluationszwecken strukturiert wird, werden nicht angegeben. Eine Programmdarstellung lässt die Einsatzpunkte der Evaluation zuordnen. Diese Darstellung weist allerdings auch Unschärfen auf und erzeugt nicht in jeder Hinsicht Klarheit, welche Zielsetzungen das Programm durch den Einsatz welcher Mittel in welcher zeitlichen bzw. logisch-hierarchischen Abstufung erreichen möchte. Es wird im Verlauf der präsentierten Analyse auch nicht zwischen direkten und indirekten Wirkungsweisen systematisch differenziert. Unter anderem wird an zwei Berichtsstellen gesagt, dass Zielerreichungen bereits durch die Entsprechung von FördernehmerInnen zu Antragsbedingungen gegeben seien, während es sich hier doch sachlich-logisch nur um eine Vorbedingung zur Erreichung von intendiertem Output, späteren Outcomes, und nachfolgenden Effekten im Sinne der Zielsetzungen handeln kann. Die Erstellung einer Logic Chart oder eine andere Form der strukturierten Aufarbeitung der Programmstruktur bzw. der logisch-hierarchischen Wirklogik unter Berücksichtigung aller existierenden Programmkomponenten hätte die Analyse jedenfalls unterstützen können. Dadurch hätte ein klar strukturierter, trennscharfer Umgang mit Zielsetzungen und darauf bezogenen Programmelementen ermöglicht werden können, der auch auf seine Vollständigkeit und Konsistenz hin gut transparent wird. Monitoringdaten werden lediglich getrennt für die beiden gemeinsam analysierten, thematisch eng zusammenhängenden Programme aufgearbeitet, es wird weder ein Gesamtbild erzeugt noch etwaigen Gründen für unter Umständen divergierende	2	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Nachfragestrukturen nachgegangen. Für vertiefte Einsichten in die Art der Wirksamkeit der Programme bzw. potenzielle Hinderungsfaktoren hätte es vielversprechend sein können, auch bisher nicht erfasste Teile der Zielgruppe und der Kontextorganisationen zu beleuchten. Dies erfolgt jedoch nicht. Unter anderem hätten dadurch Aspekte wie die Bekanntheit und Wahrnehmung des Programms tragfähiger analysiert werden können. Mit der Heranziehung nicht-erfolgreicher AntragstellerInnen als Vergleichsgruppe für die FördernehmerInnen wird eine Überprüfung von essentiellen Programmwirkungen möglich. Allerdings sind innerhalb dieser Herangehensweise die Möglichkeiten doch grundsätzlich eingeschränkt, den potenziellen Einfluss von verschiedenen relevanten Faktoren für die Entfaltung der Programmwirkungen insgesamt zu erkennen. Einige Outputs und Outcomes des Programms hätten anhand zusätzlicher objektiver Daten genauer untersucht werden können, über Selbsteinschätzungen der unmittelbar Betroffenen hinaus.	
N5 Transparenz von Werten	Es finden sich keine expliziten Darstellungen zu den Bewertungsmaßstäben, die zur Einschätzung des Programms herangezogen werden. Ein wesentlicher Orientierungspunkt der vorliegenden Analyse sind Zielsetzungen des Programms. Im Verhältnis zwischen verschiedenen Aussagen über die Programmkonzeption und der Wahl von Parametern zu deren Einschätzung (Indikatoren) finden sich allerdings auch Unschärfen. Es werden keine Überlegungen dargestellt, warum welche Parameter herangezogen werden, in welcher genaueren Weise einzelne Parameter aus Sicht der EvaluatorsInnen für den Programmserfolg Bedeutung haben, und inwiefern auch andere Parameter, die im Themenbereich des Programms eingespielt sind, möglich gewesen wären, aber mit ausgewiesenen Gründen nicht herangezogen wurden. Es ist auch nicht erkennbar, dass Kriterien vorab festgelegt worden wären, anhand derer später erhaltene Ergebnisse eingeordnet werden sollten, um dem Programm größeren oder geringeren Wert zuzumessen. Die Einschätzungen beruhen auch auf Sichtweisen von Betroffenen (FördernehmerInnen, AntragstellerInnen, zentrale Akteure im Kontext, in dem das Programm wirken will), die von den EvaluatorsInnen allerdings teils direkt übernommen werden, ohne zusätzliche Überlegungen anzustellen. Die Einschätzungen beruhen des Weiteren auf Sichtweisen von internationalen ExpertInnen, die von den EvaluatorsInnen teils direkt übernommen und teils abgelehnt werden. Die Einschätzungen beruhen schließlich auf Sichtweisen und Annahmen der EvaluatorsInnen, die an mehreren Stellen Bewertungen der Programmanlage aussprechen, ohne dass nähere begründende Ausführungen gemacht werden.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in klarer, allgemein verständlicher Sprache abgefasst und bietet zahlreiche Ergebnisse in Form von Tabellen und Grafiken an. Ein Tabellen- und Grafikverzeichnis, anhand derer diese Informationen rasch gezielt auffindbar werden, ist allerdings nicht beigegeben. Ein Executive Summary ist nicht beigegeben. Die untersuchten Programme werden in ihrer Motivation, ihrer historischen Entwicklung, ihrer Anlage sowie Veränderungen während der Laufzeit dargestellt. Dadurch werden auch bislang mit ihnen nicht vertraute BerichtsadressatInnen in die Lage versetzt, den Evaluationsgegenstand in wesentlichen Zügen zu verstehen und der Berichterstattung der Evaluation zu folgen. Allerdings ist die Vorstellung der Programme, die sowohl eine tabellarische Darstellung als auch Textpassagen umfasst, nicht hinsichtlich aller Zielsetzungen und hinsichtlich aller Komponenten der Programme, durch die deren Zielsetzungen erreicht werden sollen, klar (vgl. auch N4 und G3). Es finden sich auch rein textliche Darstellungen von Ergebnissen, die ebenfalls durch Tabellen oder Grafiken unterstützt werden hätten können. In der textlichen Berichterstattung werden nicht zu allen Ergebnissen die Gesamtverteilungen greifbar. Zumindest in einem Fall erscheint eine Tabelle unvollständig beschriftet.	2

N8 Nutzung und Nutzen der Evaluation	Aus Angaben zur Einbeziehung von Beteiligten und Betroffenen kann geschlossen werden, dass mit den HauptadressatInnen (AuftraggeberInnen) zumindest an einem Punkt im Verlauf der Evaluationsdurchführung so interagiert wurde, dass angenommen werden kann, dass die Evaluation ihren Informationsbedürfnissen gut entgegenkommt. Darüber hinaus wurden VertreterInnen der Kontext-Organisationen auf die Evaluation aufmerksam gemacht, sodass sie an ihren Ergebnissen Interesse entwickeln konnten. Während die Evaluationszwecke sowohl lernorientierte als auch an Rechenschaftslegung orientierte Komponenten enthalten, erbringt die Evaluation neben Darstellungen der Programmwirkung auch einige Ergebnisse, die von den Programmverantwortlichen für die Weiterentwicklung des untersuchten Programms hilfreich sein können. Hinweise auf konkrete Vorstellungen über genauere Nutzungsweisen oder Schritte zu ihrer Förderung und Unterstützung finden sich nicht. Die Nützlichkeit muss insgesamt wegen nicht vollkommen erfüllter Nützlichkeits-, Fairness- und Genauigkeitsstandards doch als eingeschränkt gelten.	3
F3 Vollständige und faire Überprüfung	Der Bericht ist in neutralen Formulierungen abgefasst. Er macht Stärken und Schwächen des Programms sichtbar, die allerdings nicht immer klar als solche explizit gemacht werden. Die Evaluation bringt die Erfahrungen und Sichtweisen von Betroffenen (erfolgreiche und nicht erfolgreiche Teile der Zielgruppe) sowie von wesentlichen Akteuren im Programmumfeld zur Geltung. Aussagen der Befragten werden in einigen Abschnitten durch Zitate gut nachvollziehbar, jedoch nicht durchgehend. Qualitative Ergebnisse werden teilweise so berichtet, dass nicht greifbar wird, welcher Stellenwert den berichteten Aussagen insgesamt unter allen erhaltenen Aussagen zukommt. Nicht zu allen quantitativen Ergebnissen werden Gesamtverteilungen dargestellt, sodass allein anhand der berichteten Teilergebnisse deren effektive Bedeutung von BerichtsadressatInnen nicht eingeschätzt werden kann. Nicht alle potenziellen Ansatzpunkte für die Reflexion von Gestaltungsaspekten des Programms, die anhand von berichteten Ergebnissen greifbar werden, werden explizit herausgearbeitet. In der Interpretation von Ergebnissen werden stellenweise alternative mögliche Begründungen für bestimmte Faktenlagen erläutert. Es finden sich aber auch Stellen, an denen die EvaluatorInnen Argumentationen einbringen, die zwar durchaus als eine mögliche und plausible Begründung für die Faktenlage zu erachten sind, aber ohne dass explizit reflektiert würde, ob auch andere Gründe vorliegen könnten bzw. den Ergebnissen eine andere Bedeutung zukommen könnte. Belege dafür, dass die herangezogene Argumentationsweise stichhaltig ist, werden hier nicht angegeben. Ergebnisse einer eingesetzten Methode werden nicht interpretiert und nicht in die Berichterstattung integriert, sondern nur in einem Anhang beigegeben. Bei der Präsentation von zentralen Ergebnissen und Empfehlungen für die Weiterentwicklung des Programms werden erhobene Sichtweisen von vom Programm bereits erreichten Teilen der Zielgruppe von den EvaluatorInnen direkt übernommen, ohne dass sie zu erkennen geben, inwiefern sie diese Sichtweisen – auch unter Heranziehung anderer Informationen – teilen oder zu anderen Einschätzungen gelangen.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Hinsichtlich des institutionellen Kontexts wurden die Erfahrungen und Sichtweisen wesentlicher Akteure in Erfahrung gebracht. Durch diese Vorgehensweise werden an verschiedenen Stellen Rahmenfaktoren und Mechanismen erkennbar, die die Entfaltung von intendierten Programmeffekten beeinflussen. Dabei wird auch	3

	eine innere Differenzierung dieses Kontexts nach einigen Strukturmerkmalen verfolgt. Letztlich bleiben die erbrachten Hinweise allerdings doch punktuell und es hätte noch eingehender verfolgt werden können, wie das Programm in der Interaktion mit seinem – unter Umständen auch noch in anderer als der untersuchten Weise in sich differenzierten - Kontext seine Wirkungen entfaltet bzw. an der Entfaltung intendierter Wirkungen gehindert ist. Strukturbedingungen werden letztlich nur ansatzweise reflektiert und teilweise unter Heranziehung von Plausibilitätsargumenten behandelt. Es wird berichtet, dass die untersuchten Programme neben Effekten bei den direkt Begünstigten (antragsberechtigte Zielgruppe) auch Effekte in ihrem Kontext auslösen sollen. Während eine angegebene Methode darauf schließen lässt, dass eine Untersuchung von ausgelösten Veränderungen in den Kontextorganisationen geplant war, wird darüber nichts berichtet. Darüber hinaus hätten hier theoretisch hier auch noch weitere, auf andere zusätzliche Daten gestützte Analysen stattfinden können. Der sozioökonomische Kontext der FördernehmerInnen bzw. Zielgruppe, auf den das Programm ebenso wie auf institutionelle Strukturen gezielt reagieren will, wurde nicht untersucht, sodass nicht sichtbar werden kann, ob und inwiefern hier in der aktuellen Situation Anpassungen angeraten erscheinen könnten. Andere Programme mit ähnlicher Zielsetzung oder potenziell ähnlichen Wirkungen werden in die Analyse einbezogen. Dies erfolgt allerdings nur auf der Ebene der Wahrnehmung dieser Programme durch die Zielgruppe und nicht in Form einer systematischen Analyse anhand objektiver Daten. In Entsprechung zu einem angegebenen Evaluationszweck (vgl. N2) hätte hier eine tiefergehende Analyse stattfinden können.	
G3 Beschreibung von Zwecken und Vorgehen	Die Vorgehensweise der Evaluation in methodischer und organisatorischer Hinsicht wird dargestellt, wobei auch eine Abänderung des ursprünglich geplanten Vorgehens der Evaluation dargelegt wird. Es wird nachvollziehbar, inwiefern die Evaluation von der Auftragserweiterung profitieren konnte. Die eingesetzten Methoden werden im Zusammenhang mit der Berichterstattung über die jeweils erbrachten Ergebnisse und in technischen Anhängen so dargestellt, dass deren Tragfähigkeit und Aussagekraft weitestgehend eingeschätzt und Vor- bzw. etwaige Nachteile erkannt werden können. Die genauen Fragestellungen der Umfrageerhebung werden in einigen Ergebnispräsentationen angegeben. Vollständige Fragebögen oder Interviewleitfäden sind allerdings nicht beigegeben, was die Einschätzung aller Ergebnisse und ihrer Interpretationen erst vollständig ermöglicht hätte. Es sind unterschiedliche Vorgehensweisen im Umgang mit verschiedenen Datenquellen festzustellen, wofür keine Gründe angegeben werden. Mangels einer Differenzierung von genaueren Evaluationsfragestellungen und einer systematisch strukturierten, vollständig-konzisen Programmdarstellung (vgl. N4) entzieht sich zumindest für BerichtsadressatInnen, die mit den Programmen nicht sehr gut vertraut sind, trotzdem ein Stück weit, wie vollständig und insgesamt tragfähig es untersucht wurde. Eine übersichtliche explizite Darstellung, ob und inwiefern Einschränkungen für die Aussagekraft und Tragfähigkeit der Ergebnisse vorliegen könnten, wird den BerichtsadressatInnen nicht gegeben.	2
G8 Begründete Schlussfolgerungen	Die EvaluatorInnen präsentieren in einem eigenen, von Ergebnisdarstellungen klar getrennten Kapitel ihre Schlussfolgerungen. Beobachtete Fakten werden dabei ausgewiesen und klar eingeordnet. In den Schlussfolgerungen wird auf eine Darstellung von Programmzielsetzungen Bezug genommen, die einige an anderen Stellen benannte mittel- und längerfristige, übergeordnete oder indirekte Zielsetzungen, insbesondere in Bezug auf den Programmkontext (vgl. G2), nicht umfasst, sodass zu solchen Zielen keine Schlussfolgerungen vorliegen. Teile der Argumentation werden unter Bezug auf ein Programmmotiv vorgetragen, von dem in der vorangegangenen Untersuchung unklar bleibt, ob es sich um eine ursprüngliche	2

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Zielsetzung des untersuchten Programms handelt, welcher Stellenwert diesem Motiv in der Entfaltung von direkten und indirekten Wirkungen zukommen sollte, und welche Mittel zur Erreichung eingesetzt werden sollten. Verschiedene Abschnitte der Schlussfolgerungen sind diesbezüglich nicht völlig konsistent. Einige Empfehlungen werden ausgesprochen, die teils spezifisch und in der Programmgestaltung umsetzbar sind, teils aber auch nur skizzenhafte Hinweise auf Gestaltungsoptionen darstellen. Es finden sich hier auch Benennungen von Ideen, zu denen die EvaluatorsInnen nicht angeben, ob und warum sie sie befürworten oder ablehnen. Zusätzlich gibt ein Kapitel Empfehlungen für die Weiterentwicklung des Programms, die primär an die umsetzende Agentur adressiert sind. Hier besteht enger und transparenter Bezug zu Analyseergebnissen.	
--	--	--

Evaluationsbericht Nr. 13		Beobachtungszeitraum: 2007 - 2010	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In Datenerhebungsverfahren wurden für das Programm Verantwortliche, mit seiner Umsetzung Befasste sowie wesentliche Akteure der vom Programm betroffenen FTI-Landschaft wurden einbezogen. Des Weiteren wurden hier Akteure, die informierte BeobachterInnen darstellten oder auf Grund ihrer Positionen am Programm und an Erkenntnissen darüber Interesse haben konnten, einbezogen. Für das Programm Verantwortliche und mit seiner Umsetzung Befasste wurden in Diskussionen zu Methodendurchführung, Zwischenergebnissen und voraussichtlichen Empfehlungen einbezogen. Aus Angaben zum Evaluationsprozess erschließt sich, dass eine Interaktion zwischen EvaluatorInnen und AuftraggeberInnen zumindest an 5 vorab formalisierten Zeitpunkten erfolgte. Die dargestellte Interaktionsdichte kann als ausreichend dafür gelten, die Informationsbedürfnisse des Auftraggebers und etwaige Veränderungen darin zu erkennen.	1	
N2 Klärung der Evaluationszwecke	Angegeben wird: Reflexion des bisherigen Programmverlaufs, um auf dieser Basis Schlussfolgerungen und Empfehlungen für die Weiterentwicklung des Programms im Gesamtkontext des österreichischen Förderungssystems zu entwickeln. Good Practice sowie Ansätze für Verbesserungspotenziale, Adaptionen oder eine veränderte Ausrichtung hinsichtlich der Steuerung und des Designs identifizieren. Die Positionierung des Programms in der österreichischen Förderungslandschaft bestimmen. Ein abstrakter Bezug zu den Evaluationsvorgaben in den Programmrichtlinien wird hergestellt. Die angegebenen Untersuchungsschwerpunkte sind recht breit, ohne dass eine Priorisierung bestimmter Gesichtspunkte zu erkennen wäre.	3	
N4 Auswahl und Umfang der Informationen	Detaillierte Evaluationsfragestellungen werden gelistet. In der Formulierung der Evaluationsfragestellungen wird der in der Programmkonzeption angenommene Zusammenhang zwischen Zielen, Maßnahmendesign und intendierten Wirkungen adressiert, sowie die Konsistenz und Kohärenz der gewählten Instrumente in Bezug auf die Programmzielsetzungen. Des Weiteren wird auf eine Identifikation von hinderlichen intervenierenden Faktoren hingearbeitet. Mehrere Sub-Fragestellungen haben mit einer Verfolgung von Wirkungslogiken bzw. Voraussetzungen für das Eintreten des erwarteten Programmerfolgs zu tun. Die Evaluation stützt sich auf den Einsatz von vier Methoden, mit qualitativer Schwerpunktsetzung. Die eingesetzten Methoden erscheinen geeignet, die aufgeworfenen Evaluationsfragen abzudecken und zu beantworten. Ergebnisse früherer Evaluationen des untersuchten Programms und des relevanten Programmportfolios der umsetzenden Agentur werden aufgegriffen. Elemente einer Wirklogik und Voraussetzungen, unter denen das Erreichen von gesetzten Zielen als wahrscheinlich erachtet werden kann, werden jedoch letztlich nur im Stil einer Review benannt. Während hier verschiedene Aspekte der Programmgestaltung und -umsetzung sowie des Programmkontextes sichtbar werden, kann doch letztlich nicht davon gesprochen werden, dass die Wirklogik gesamthaft-systematisch und tiefgehend untersucht worden wäre. Die Untersuchung gibt sich streckenweise mit Daten bzw. Analysen vorliegender Daten zufrieden, die intensiver und umfangreicher angestellt werden hätten können, um das Programm so gut wie	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	möglich zu verstehen. • In der Darstellung von Programmoutput verlässt sich die Evaluation auf innerhalb des Programms erstellte Monitoringdaten. In der Programmkonzeption optional vorgesehene, über die monetär geförderten Aktivitäten hinausgehende Aktivitäten der Fördernehmer, die für die Gesamtwirkung des Programms Stellenwert haben könnten, werden nicht untersucht. • Das Programm will seine Wirkungen in einer zweistufigen Logik gegenüber direkt Begünstigten (unmittelbare Leistungsempfänger) und indirekt Begünstigten erreichen. Dieser wesentliche konzeptive Gesichtspunkt wird nicht gut aufgegriffen, da die Evaluation keine detaillierte Analyse der vorgesehen Programmaktivitäten auf allen Ebenen durchführt. • Partner-Netzwerke der direkten Fördernehmer, die auf die Entfaltung der vom Programm angesprochenen Aktivitäten und deren Effekte einen Einfluss haben könnten, werden kurz angesprochen, aber nicht gezielt untersucht. • Hinsichtlich der direkten Förderung werden Förderauflagen erwähnt, aber nicht als Bedingungen der Wirkungsweise der geförderten Aktivitäten aufgegriffen. • Zentrale Fragestellungen zur Anlage des Programms tauchen im Zuge von Datenauswertungen auf, die weiter zu verfolgen sich zweifellos gelohnt hätte. Hierzu kann geschlossen werden, dass eine phasenweise, flexible Gestaltung des Evaluationsprozesses helfen hätte können, diesen Erkenntnissen und ihrer weiteren Bedeutung noch intensiver nachzugehen. • In Bezug auf einen wesentlichen Aspekt der intendierten Programmwirkung verweisen die EvaluatorenInnen auf die Notwendigkeit zusätzlicher, getrennt anzustellender Untersuchungen. • Die Evaluation nimmt in ihrer Auseinandersetzung mit der Relevanz des Programms eine Bedarfseinschätzung vor, die sich auf Stellungnahmen von Betroffenen und potenziell Betroffenen des Programms stützt. Sie zieht jedoch keine objektiven Daten heran, und sie unterhält keine Beziehung zu einer im Bericht erwähnten, der Programmkonzeption zugrunde liegenden Bedarfsanalyse. • An verschiedenen Stellen wird angemerkt, dass sich die Konfiguration der geförderten Einrichtungen regional unterschiedlich gestaltet, sodass unterschiedliche Faktorenkonstellationen hinsichtlich der Erreichbarkeit der gesetzten Ziele anzunehmen sind. Diesen Konstellationen wird jedoch nicht weiter nachgegangen. Dies erscheint für das Erkennen von Faktoren, die die Wirklogik des Programms fördern oder behindern können, nachteilig. • Im angestellten Vergleich mit ähnlich gelagerten internationalen Programmen wird deren Positionierung im jeweiligen „Ökosystem“ des nationalen Förderportfolios bzw. Steuerungsinstrumentariums gegenüber dem Innovationssystem nicht tiefergehend erfasst. Dies hätte deutlich umfangreicher Untersuchungen bedurft. Es kann davon ausgegangen werden, dass dieses Erkenntnisinteresse im Rahmen der breiten Zwecksetzungen der Evaluation durch andere, gleichzeitig anwesende Untersuchungsschwerpunkte konterkariert wurde. Die Einschätzung des untersuchten Programms hat einen expliziten Bezugspunkt im international etablierten Konzept der „Knowledge Regions“, das unter Bezugnahme auf einige internationale Quellen dargestellt wird. Die vorfindliche Analyse hat letztlich allerdings doch wenig Konnex zu diesem Bezugsrahmen, es tauchen auch Argumente aus dem Konzeptrahmen der klassischen Ökonomie auf.	
N5 Transparenz von Werten	Eine explizite Darlegung von Überlegungen und Entscheidungen, wie dem Programm Wert zugemessen werden soll, findet sich nicht. Bezugspunkt der Einschätzungen sind die Zielsetzungen des Programms. Die Darstellung lässt mangels Quellennachweisen zu den dargestellten Zielen allerdings keine Rückschlüsse darauf zu, inwiefern die behandelten Zielsetzungen der ursprünglichen Programmkonzeption entstammen, und inwiefern sie eventuell im	3

	Zuge der Erarbeitung der Evaluation teilweise abweichend davon formuliert wurden, z.B. um einem aktuellen politischen Interesse entgegen zu kommen. Sichtweisen auf das Programm aus unterschiedlichen Akteursgruppen werden zu einigen, jedoch nicht allen, Analysepunkten herangezogen. Es wird dabei auch über Differenzen zwischen Sichtweisen unterschiedlicher Akteursgruppen berichtet, wenn auch nicht durchgehend. In der gewählten Analysestrategie tritt tendenziell eine Dominanz der Perspektive der Fördergeber ein, was mit der Auswahl der herangezogenen Information in direktem Zusammenhang zu sehen ist. Dies äußert sich nicht zuletzt darin, dass Effekte auf die indirekten Begünstigten, die für die Einschätzung des Programms eigentlich auf Grund seiner Anlage als wesentlich erachtet werden sollten, in der vorgenommenen Analyse nur eine untergeordnete Rolle spielen. Mit dem bereits genannten Changieren zwischen zwei Konzepten zur Einordnung der Beobachtungen (vgl. N4) ergeben sich Unsicherheiten für die Gesamteinschätzung.	
N6 Vollständigkeit und Klarheit der Berichterstattung	Die Berichtssprache ist klar und von Technizismen frei. Der Bericht ist klar gegliedert und gut handhabbar. Er bietet Informationen auf Konzept- und Ergebnisebenen in grafischer und Tabellenform an. Unterschiedliche Schriftfarben werden eingesetzt, um die Handhabung und Auffassbarkeit zu erhöhen. Ein einseitiges Executive Summary stellt übersichtlich das Programm mit seiner Zielsetzung und Eckparametern seines Operierens, die Hauptfragestellungen der Evaluation einschließlich der Voraussetzungen der Evaluierbarkeit, Ergebnisse zu Programmoutputs und -outcomes, sowie Empfehlungen dar. Auf das Programm wird hinsichtlich seiner Motivation, Geschichte, Gestaltungsweise und Operationsweise eingegangen, sodass der Evaluation auch durch nicht mit dem Programm Vertraute gefolgt werden kann. Ergebnisse werden so dargestellt, dass die BerichtsadressatInnen verstehen können, welchen Stellenwert die EvaluatorenInnen ihnen beimessen.	1
N8 Nutzung und Nutzen der Evaluation	Auf Grund der mehrfachen Interaktionen zwischen EvaluatorenInnen und AuftraggeberInnen im Evaluationsprozess (vgl. N1) kann davon ausgegangen werden, dass die Evaluation gute Voraussetzungen dafür geschaffen hat, für ihre HauptadressatInnen nützlich zu sein. Gesichtspunkte der Präsentation und der methodischen Transparenz lassen die Diagnose zu, dass ein Evaluationsbericht vorliegt, der von seinen AdressatInnen gut genutzt werden kann. Daneben wurden wesentliche Akteure der vom Programm betroffenen FTI-Landschaft auf die Evaluation aufmerksam gemacht. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden soll, werden allerdings nicht dargestellt. Einschränkungen für die Nützlichkeit ergeben sich aus einer nur bedingten Erfüllung einiger Nützlichkeits-, Fairness- und Genauigkeitsstandards.	3
F3 Vollständige und faire Überprüfung	Der Bericht ist durchgängig in einer neutralen, faktenorientierten Sprache abgefasst. Der Bericht präsentiert Stärken und Schwächen des untersuchten Programms. Das Schlusskapitel zu Schlussfolgerungen und Empfehlungen gliedert klar zwischen identifizierten Stärken und Schwächen. Der Bericht ist jedoch bei der expliziten Benennung von Schwächen, aus denen in der Folge gelernt werden könnte, wie es den angegebenen Evaluationszwecken entspricht, eher zurückhaltend. Dadurch wird die Auffassbarkeit dieses Lernpotenzials für Adressatenkreise jenseits von unmittelbar mit dem Programm Vertrauten tendenziell geschwächt. Manche identifizierte Schwächen werden in Resümees aufgegriffen, jedoch dann außerhalb des Verantwortungsbereichs des Programms gestellt. Nicht alle anklingende Aspekte, die unter Umständen	3

	Hinderungsgründe für die Erreichung von Programmserfolg im Rahmen der bestehenden Programmkonzeption dargestellt haben könnten, werden auch konsequent zum Ausdruck gebracht. Sichtweisen von Beteiligten oder Betroffenen fließen in die Analyse ein. Diese werden allerdings teilweise direkt übernommen und teilweise nicht aufgegriffen, wobei nicht an allen Stellen transparente Argumentationen oder Faktenverweise vorgelegt werden. Sichtweisen von internationalen ExpertInnen mehrerer Länder wurden in einem Workshop zu den vorläufigen Evaluationsergebnissen einbezogen. Allerdings werden die Ergebnisse dieses Workshops getrennt dargestellt und es bleibt unklar, inwiefern die EvaluatorenInnen diese ExpertInnensichten aufgegriffen oder mit guten Gründen nicht integriert haben.	
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das Förderangebot bzw. -geschehen wird auch auf mögliche Synergien, Überschneidungen und Anknüpfungspunkte zu benachbarten Programmen hin untersucht. Dies beschränkt sich allerdings auf relevante andere Programme auf Bundesebene, Programme auf Länderebene werden nicht betrachtet. In den gezogenen Schlussfolgerungen werden Verbesserungsmöglichkeiten auch außerhalb der Reichweite des Programms in einem anderen Steuerungsbereich des FTI-Systems benannt. Hinsichtlich der Positionierung im Förderportfolio des Bundes begnügt sich die Evaluation letztlich mit einer vergleichenden Zusammenstellung der Positionierung der verschiedenen Programme von ihrer Anlage her. Eine eingehende Analyse des ausgelösten Fördergeschehens und seiner Effekte im Kontext des Portfolios, aus dem sich der Beitrag des untersuchten Programms erst substanziell erschließen ließe, wird nicht präsentiert. Die Programmkonzeption wird in den Kontext internationaler Entwicklungen und Herausforderungen gestellt, in dem es sich mit seiner Konzeption ansiedelt. Die Evaluation untersucht diesbezüglich auch die Konzeption ausgewählter internationaler Vergleichsprogramme. Hier wird eine Systematisierung angeboten, die allerdings hinsichtlich der faktischen Wirklogiken und der Rollen der Programme im Verhältnis zu sie umgebenden Strukturen und anderen Initiativen kaum Aussagen enthält, die für ein weiteres Lernen über das untersuchte Programm verwertbar wären. Hier ist jedenfalls zuzugestehen, dass diese Fragestellung von hoher Komplexität gekennzeichnet ist und zweifellos umfangreicher Untersuchungen bedurft hätte. Hinsichtlich der Identifikation von Good Practice, die in den Evaluationszwecken genannt ist, besteht in der internationalen Evaluationstheorie die Auffassung, dass dazu das untersuchte Programm möglichst genau auf die Verknüpfung mit seinem Kontext untersucht werden sollte. Dem konnte die vorliegende Analyse ein Stück weit gerecht werden, hätte diesbezüglich aber zweifellos auch nicht vertieft werden können.	3
G3 Beschreibung von Zwecken und Vorgehen	Die Ziele der Evaluation werden im Eingangskapitel benannt und es werden Evaluationsfragestellungen gelistet, gegliedert in Haupt- und Sub-Fragestellungen. Die gelisteten Evaluationsfragestellungen sind geeignet, den angegebenen Evaluationszweck bzw. die Zielsetzung der Evaluation zu erfüllen. Die Subfragen erscheinen gegenüber den Hauptfragestellungen stringent konzeptualisiert. Ein eigenes Teilkapitel geht auf Voraussetzungen der Evaluierbarkeit ein, ein Weiteres sodann auf die methodische Vorgehensweise der Evaluation. Es ist gut nachvollziehbar, wie die gewählten Methoden im konzeptuellen Aufbau der Analyse eingesetzt werden. Es besteht ein enger, transparenter Konnex zwischen den Fragestellungen der Interviews und den Evaluationsfragestellungen. Für die Ergebnisse der Dokumentenanalyse ist dies in der Form, in der über sie berichtet wird, weniger der Fall.	2

	An einigen Stellen wird darauf hingewiesen, dass verfügbare Daten für die Beantwortung von Fragestellungen nicht ausreichen. Es wird jedoch nicht systematisch, durchgehend und leicht auffindbar berichtet, welche Aspekte des Evaluationskonzepts mithilfe der eingesetzten Methoden wie gut eingelöst werden konnten, und wo die Beantwortung der gelisteten Fragestellungen eingeschränkt oder weniger tragfähig ist. Die Ergebnisinterpretation ist nicht mit einer Auseinandersetzung versehen, dass die Einschätzung der Erreichung von Zielen durch Programmbeteiligte mit einem Bias behaftet sein könnte. Angaben zu den eingesetzten Methoden, die deren etwaige Vor- und Nachteile und die Tragfähigkeit der Ergebnisse einschätzen lassen, sind für manche, aber nicht alle Methoden vorhanden. Es wird nicht bündig dargelegt, ob und inwiefern Einschränkungen in der Behandlung vorgesehener Untersuchungsschwerpunkte und Evaluationsfragestellungen bestehen könnten.	
G8 Begründete Schlussfolgerungen	Schlussfolgerungen und Handlungsempfehlungen werden in einem eigenen, von Analyseschritten und Dateninterpretationen getrennten Kapitel dargestellt. Eine innere Strukturierung des betreffenden Kapitels gewährleistet eine transparente Differenzierung der Argumente. Schlussfolgernde Passagen finden sich auch im Berichtstext, wo sie in der Regel auch klar markiert und demzufolge gut auffindbar sind. Die Schlussfolgerungen und Empfehlungen heben weitestgehend gut erkennbaren weitestgehend Rückhalt in den präsentierten Daten. Es finden sich allerdings auch punktuell Schlussfolgerungen, die allgemein-abstrakten Aussagen machen, die nicht auf erbrachten Ergebnissen zum Evaluationsgegenstand aufbauen. Stellenweise werden nicht alle berichteten Fakten, mit denen die hinter den Schlussfolgerungen stehenden Fragen beantwortet werden, vollständig aufgegriffen. An mehreren Stellen wird mit einer bestimmten Wirkungsweise des Programms argumentiert, während das Operieren des Programms in dieser Hinsicht nicht eingehend untersucht wurde.	3

Evaluationsbericht Nr. 14		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	summativ (Endbericht einer Begleitevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Der Bericht, der eine abschließende Analyse vorlegt, nachdem das untersuchte Programm bereits zuvor Zwischenevaluationen unterzogen wurde, enthält keine Angaben, ob und inwiefern Beteiligte oder Betroffene in die Planung und Durchführung der Evaluation einbezogen wurden. Aus getrennten Berichten zu den Vorgängerevaluationen ist zu entnehmen, dass an der Programmumsetzung Beteiligte und VertreterInnen der Zielgruppe im Rahmen von Datenerhebungsverfahren einbezogen wurden. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen weitere FTI-politische Akteure einbezogen worden wären, die an der Evaluation und ihren Ergebnissen ein Interesse haben könnten.	4	
N2 Klärung der Evaluationszwecke	Es wird angegeben: Aufbauend auf vorliegenden Berichten aus dem vorangegangenen Begleit- und Monitoringprozesses über die gesamte Programmlaufzeit wesentliche Lehren und Schlussfolgerungen ziehen. Die wesentlichsten Lehren aus existierenden Evaluationsergebnissen zusammenfassen und der Frage nachgehen, wie die öffentliche Hand nach Programmende im Operationsbereich des untersuchten Programms weiter tätig werden kann, und wie eine etwaige Neuausrichtung von Fördermaßnahmen aussehen könnte. Die Evaluationszwecke werden somit in sehr breiter und unscharfer Weise umrissen. Aus der Formulierung, die auf die zukünftige Gestaltung von Fördermaßnahmen Bezug nimmt, kann auf eine Orientierung an konzeptuellem Nutzen geschlossen werden. Mit der Benennung von einigen Evaluationsthemen wird die Evaluationsaufgabe konkretisiert und eingegrenzt. Es ist allerdings nicht zu entnehmen, inwiefern diese Strukturierung abgeklärten Informationsbedürfnissen der HauptadressatInnen (AuftraggeberInnen) entsprochen hat.	4	
N4 Auswahl und Umfang der Informationen	Ausgehend vom groben Umriss der Evaluationszwecke werden drei Themenfelder benannt, auf die sich die Evaluation konzentriert. Zu deren Behandlung werden drei Methoden parallel eingesetzt, die als solche für die Analyse gut geeignet sind. Auf das Programm wird hinsichtlich von Motivation, Zielsetzungen und Anlage detailliert eingegangen. Die verschiedenen Schritte und Komponenten, durch die das Programm die Erreichung seiner Ziele verfolgt, werden aufgeschlüsselt. Diese Aufschlüsselung kommt ohne den Einsatz einer formalen Logic Chart aus. Die Aufschlüsselung der Programmlogik wird allerdings nicht in nachvollziehbarer Weise genutzt, um in der weiteren Analyse Elementen der Programmgestaltung hinsichtlich ihres Stellenwerts für die Erreichung von Outputs, Outcomes und Wirkungen nachzugehen. Die Relevanz des Programms wird anhand von internationaler Fachliteratur behandelt. In dieser Diskussion werden grundlegende Mechanismen des Innovationsgeschehens und der Verhaltensmuster von Akteursgruppen in diesem Geschehen benannt. In der weiteren Berichterstattung wird jedoch diesen Faktoren nicht im konkreten Kontext weiter nachgegangen. Hinsichtlich eines Weiterbestehens der Relevanz zum Evaluationszeitpunkt werden keine geeigneten Analysen angestellt (z.B. Sichtweisen der Betroffenen, Bedarfserhebung). Die Evaluation hat über weite Strecken den Charakter eines Wirkungsmonitorings, wobei über erhaltene Monitoringdaten nicht durch eigene, zusätzliche Datengewinnungen hinausgegangen wird.	4	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Programmoutcomes werden in einer Weise untersucht, die deren Entwicklung während der Programmlaufzeit aufzeigt. Es werden allerdings nicht Veränderungen gegenüber einer vor dem Start des Programms liegenden Periode untersucht (Pre-Post-Design), was klarere Erkenntnisse über ausgelöste Effekte erbringen hätte können. Hier ist zuzugestehen, dass dies nur bei frühzeitiger Planung und mit einem erheblichen Mehraufwand zu bewerkstelligen gewesen wäre. Es wird auch keine Relation zwischen Outcomes und Umsetzungsschritten im untersuchten Programm hergestellt, die das Auftreten dieser Outcomes beeinflusst haben könnten. Es wird auch nicht der Frage nachgegangen, ob andere Einflussgrößen außer dem untersuchten Programm mitverantwortlich sein könnten, sodass alle festgestellten Outcomes dem untersuchten Programm zugerechnet werden. Die Erreichung derjenigen Programmziele, die in der Programmdarstellung genannt sind, wird teils oberflächlich und teils nicht in geeigneter Form überprüft. Mögliche Erkenntnisse aus vorhandene Daten werden nicht gezielt aufbereitet. Darüber hinaus wären zusätzliche Daten aus ergänzenden Methoden benötigt worden, um das Ausmaß von Zielerreichungen gut feststellen zu können. Dies hätte auch die Voraussetzung dargestellt, um sich Aspekten der Programmumsetzung sowie Randbedingungen (intervenierenden Faktoren) zuzuwenden, die das beobachtete Ausmaß von Zielerreichungen beeinflusst haben können. Entwicklungen, die im organisationellen Kontext im Sinne der Programmzielsetzungen eingetreten sind, hätten durch ergänzende Datenerhebungen beleuchtet werden können (vgl. auch G2). Es wird unter anderem eine Fragestellung untersucht, die in der dargestellten Programmanlage keine klare Motivation hat. Die Evaluation konzentriert sich somit nicht durchgehend auf die möglichst gute Behandlung der wesentlichsten Gesichtspunkte, wie sie unter Bezugnahme auf die ausgewiesenen Evaluationszwecke und -themen aufzufassen sind.	
N5 Transparenz von Werten	Es finden sich keine Angaben über Bewertungsmaßstäbe, die zur Einschätzung des Programms herangezogen werden. Vorgenommene Einschätzungen stützen sich auf die Sichtweisen der EvaluatorInnen, die sie aus ihrem disziplinären Hintergrund heraus in der Beobachtung von objektiven Daten entwickeln. Sichtweisen von Beteiligten und Betroffenen des untersuchten Programms werden nicht herangezogen. Alternative Einschätzungen, die sich aus der Perspektive von Programmumsetzenden oder Der Diskussion der Relevanz des untersuchten Programms liegt erkennbar ein innovationsökonomischer Ansatz mit seinem typischen Hauptbezugspunkt des Marktversagens zugrunde. Zielerreichungen stellen einen Orientierungspunkt dar, werden aber nur teilweise in einer Weise überprüft, wie es zur berichteten Programmanlage als angebracht zu erachten wäre. Vorab festgelegte Kriterien, die zur späteren Einstufung von erhaltenen Ergebnisse dienen, werden nicht erkennbar. In der Untersuchung von Outcomes wird nicht transparent, auf welcher Basis dem Programm hier Wert zugewiesen werden soll, und die EvaluatorInnen enthalten sich einer expliziten Feststellung von Stärken oder Schwächen. Es wird offenbar wie selbstverständlich davon ausgegangen, dass die berichteten Daten hinsichtlich von Marktentstehung bzw. Marktteilnahme aus sich selbst sprechen. In den Schlussfolgerungen werden auch Bewertungen vorgenommen, deren Grundlage nicht explizit gemacht wird.	4
N6 Vollständigkeit und Klarheit der Berichterstattung	Der übersichtliche und klar strukturierte Bericht ist in allgemeinverständlicher Sprache abgefasst. Zentrale Ergebnisse werden mittels Tabellen und Grafiken aufbereitet. Ein Tabellen- und Grafikverzeichnis, mithilfe dessen sie leicht aufgefunden werden können, wurde allerdings nicht beigegeben. Dem Bericht ist ein kurzes Abstract im Umfang eines Absatzes vorangestellt. BerichtsadressatInnen müssen daher für mehr als diese recht abstrakten Informationen in den Berichtstext	2

	hineingreifen. Dies erscheint angesichts der Kürze des vorliegenden Berichts eher unproblematisch. Das untersuchte Programm wird eingehend dargestellt, sodass auch mit ihm bislang nicht Vertraute der Berichterstattung gut folgen können. Auch Hinweise auf den Programmkontext sind ausreichend vorhanden. Einige Programmeffekte werden anhand einer nachvollziehbar gemachten Methode identifiziert. Andere Effekte werden darüber hinaus in einer pauschalisierenden Weise berichtet, ohne dass dahinter stehende Methoden und Ergebnisse transparent gemacht würden. Sofern es sich hier um Übernahmen von Ergebnissen aus Vorgängerevaluationen handeln sollte, fehlen entsprechende Hinweise.	
N8 Nutzung und Nutzen der Evaluation	An der Programmumsetzung Beteiligte und Betroffene aus der Zielgruppe des untersuchten Programms wurden im Rahmen von Datenerhebungen in einzelnen Abschnitten des Gesamt-Evaluationsverfahrens einbezogen, sodass sie potenziell auch an den Evaluationsergebnissen Interesse entwickeln konnten. Angaben über die Einbeziehung der AuftraggeberInnen und etwaiger weiterer Programmverantwortlicher, die auf die Vermittlung von deren Informationsbedürfnissen schließen lassen könnten, werden nicht gemacht. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden sollen, werden nicht dargestellt. Sofern den angegebenen Evaluationszwecken eine Orientierung an konzeptuellem Nutzen entnommen werden kann, werden diesbezüglich nur eingeschränkt Informationen erbracht, die dem Lernen für zukünftige Maßnahmen im Themenbereich des Programms tatsächlich dienen können. Die Nützlichkeit muss angesichts niedrigerer Erfüllungsniveaus mehrerer Nützlichkeits-, Fairness- und Genauigkeitsstandards als deutlich eingeschränkt eingestuft werden.	4
F3 Vollständige und faire Überprüfung	Der Bericht ist in neutraler Sprache abgefasst. Abgesehen von der neutralen Formulierungsweise zeigen sich jedoch Nachteile hinsichtlich der Aussagekraft, zu welcher Ansicht die EvaluatorsInnen über das untersuchte Programm im Verlauf ihrer Analysen gelangt sind. Die Diskussion der Relevanz des Programms anhand internationaler Fachliteratur verhandelt in erster Linie Erwartungen, die auf Grund internationaler Erfahrungen im Operationsbereich des Programms an seine Wirkungsweise gerichtet werden können. Es werden jedoch keine konkreten Fakten aus dem tatsächlichen Operationsbereich des Programms herangezogen. Es bleibt letztlich unklar, inwiefern die benannten Wirkungsweisen von den EvaluatorsInnen für das untersuchte Programm in seinem konkreten Kontext ebenso angenommen werden. In einem Berichtsabschnitt werden dem Programm Wirkungen zugesprochen, die die EvaluatorsInnen jedoch nur beispielhaft und ohne Angaben von Belegen beschreiben. Es werden in der insgesamt deskriptiven Vorgehensweise, die primär auf Datenpräsentationen setzt, kaum Stärken oder Schwächen des Programms explizit herausgearbeitet.	4
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können	3
G2 Kontextanalyse	Das Programm wird im Kontext internationaler Entwicklungen begriffen, die in einigen relevanten Eckpunkten dargestellt werden. Internationale Erfahrungen im Themenbereich des untersuchten Programms werden herangezogen, um grundlegende Wirkungsweisen bzw. Wirkungsvoraussetzungen zu identifizieren. Dabei wird auch darauf hingewiesen, dass grundsätzliche Probleme mit der Übertragbarkeit von in einzelnen Ländern gewählten Strategien und von Erkenntnissen über die Wirksamkeit solcher Strategien bestehen. Im Rahmen der durchgängigen	4

	Herangehensweise der Evaluation, deren Argumentationsweise an allgemeinen Gesetzmäßigkeiten orientiert ist, werden punktuelle Erkenntnisse aus unterschiedlichen nationalen Kontexten zu einem verallgemeinerten Bild zusammengezogen, das kontextuelle Randbedingungen vernachlässigt. Das Programm wird im Kontext von Handlungsvoraussetzungen und Entwicklungen in den Organisationen betrachtet, auf die es den berichteten Zielsetzungen gemäß wirken will. Rechtliche Rahmensetzungen und Veränderungen in Steuerungsinstrumenten, die während der Programmlaufzeit zu weiteren synergetischen Beiträgen zu Programmwirkungen führen sollten, werden detailliert beschrieben. Allerdings werden die konkreten Vorgänge, die im Einsatz von Steuerungsinstrumenten und im autonomen Handeln der Organisationen eingetreten sind, zu denen grundsätzlich Unterlagen erhältlich gewesen wären, nicht untersucht. Wirkungen des Programms in seinem institutionellen Kontext werden beschrieben, ohne freilich konkrete Belege auszuweisen. Dass unterschiedliche Wirkungsweisen in verschiedenen Organisationen vorliegen, wird allgemein angemerkt, aber nicht genauer verfolgt. Eminente Vertreter der internationalen Evaluationsforschung weisen darauf hin, dass für die Extraktion von aussagekräftigen „lessons learned“ eine gute Beleuchtung der Kontextabhängigkeit von Programmeffekten wichtig ist. Während ein derartiger Evaluationszweck angegeben ist, wird doch dieser Anforderung in der vorliegenden Analyse kaum Genüge getan.	
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht macht keine Angaben über die zur Behandlung der Evaluationsthemen eingesetzten Methoden. Diese müssen von den Adressaten im Verlauf der Lektüre selbst erkannt werden. Auf das Programm wird hinsichtlich von Motivation und Anlage detailliert eingegangen. Allerdings werden Programmeffekte und damit möglicher Weise in Verbindung stehende Umsetzungsaspekte nicht vollständig in direktem Bezug zur dargestellten Programmanlage untersucht, ohne dass dazu eine Erläuterung gegeben würde. Bei der Analyse der Monitoringdaten werden methodische Angaben beigelegt, die auch auf Datenlücken hinweisen. Es wird allerdings nicht explizit gemacht, inwiefern daraus Einschränkungen für die Tragfähigkeit der Ergebnisse vorliegen könnten. Den BerichtsadressatInnen wird keine Darstellung angeboten, aus der sie Stärken und Schwächen bzw. Vor- und etwaige Nachteile des Vorgehens der Evaluation erschließen könnten. Es erfolgt keine Darstellung der organisatorischen Vorgehensweise der Evaluation, was insbesondere im gegebenen Zusammenhang eines stufenweisen Aufbaus aus mehreren Evaluationen von Interesse gewesen wäre. Wesentliche Angaben, die die Einschätzbarkeit der gesamten mehrschrittigen Evaluation und des vorliegenden Berichts betreffen, können nur den getrennten Berichten zu den verschiedenen Vorgängerevaluationen entnommen werden. Dies wird für einen Abschlussbericht, der die wesentlichen Ergebnisse aller Vorgängerevaluationen aufarbeiten will, als unzureichend erachtet.	5
G8 Begründete Schlussfolgerungen	Ein eigenes Kapitel stellt Schlussfolgerungen und Empfehlungen getrennt von den vorangegangenen Analysen dar. Schlussfolgerungen werden dabei nicht von Empfehlungen gut erkennbar differenziert. Hier finden sich auch Schlussfolgerungen, die anhand der zuvor berichteten Daten nicht nachvollziehbar sind. Dem Programm werden Wirkungen zugesprochen, für deren gesichertes Erkennen keine geeigneten Analyseschritte dargestellt werden. Sofern es sich hier um Übernahmen von Ergebnissen aus früheren Evaluationen des Programms handeln sollte, fehlen entsprechende Verweise, die den Faktenbezug transparent machen würden. Handlungsmöglichkeiten für die öffentliche Hand werden aus einer allgemein-prinzipiellen Argumentation heraus aufgezeigt. Die Argumentationsweise der Schlussfolgerungen bezieht sich teilweise, aber nicht immer auf konkret beobachtete Faktoren, die die tatsächliche Wirkungsweise des Programms in positiver oder	4

	negativer Weise beeinflusst haben, sodass gut gesicherte Rückschlüsse dazu möglich werden, in welcher Weise Nachfolgeprogramme oder andere Steuerungsinstrumente gestaltet werden könnten bzw. sollten. Es wird lediglich ein Set an Handlungsoptionen präsentiert. Etwaige Alternativen, die unter Umständen ebenfalls zu Verbesserungen angesichts von Schwächen oder Bedarfslagen führen könnten, werden nicht angesprochen.	
--	---	--

Evaluationsbericht Nr. 15		Beobachtungszeitraum: 2010 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluierung)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In Datenerhebungsverfahren wurden Programmverantwortliche, in bestimmten Rollen an der Programmumsetzung Beteiligte, sowie zahlreiche VertreterInnen der Zielgruppen einbezogen. In einem weiteren Schritt wurden VertreterInnen der Auftraggeber-Institutionen und Stakeholder im Umfeld des untersuchten Programms involviert. Es ist entnehmbar, dass Programmeigentümer und mit der Umsetzung Befasste in einem Kick-Off-Meeting einbezogen wurden. Weitere Hinweise auf Interaktionen mit diesen HauptadressatInnen der Evaluation während des Evaluationsprozesses finden sich nicht. Es liegt kein Hinweis vor, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben konnten.	3	
N2 Klärung der Evaluationszwecke	Es wird angegeben: Das Programm hinsichtlich Konzept, Umsetzung, Wirkung und Zielerreichung bewerten, um Empfehlungen bezüglich der Weiterführung des Programms zu formulieren. Es wird erwähnt, dass Evaluationsfragestellungen im Hintergrund dieser gesamthaften Formulierung stehen. In verschiedenen Berichtsabschnitten werden sodann explizite Fragestellungen ausgewiesen, anhand derer die Untersuchung der umfangreichen Untersuchungsschwerpunkte strukturiert ist. Es ist jedoch nicht klar, inwieweit dies die ursprünglich vorgesehenen Evaluationsfragestellungen sind, und daher auch nicht, wie klar der Evaluationsauftrag formuliert war.	2	
N4 Auswahl und Umfang der Informationen	Die Evaluation stützt sich auf den Einsatz mehrerer Methoden, die als solche und in ihrer Gesamtheit als gut geeignet zu erachten sind. In der Kombination von qualitativen und quantitativen Verfahren werden sowohl die Wirkungs- als auch die Organisations- und Umsetzungsebene beleuchtet. Auf dieser Basis gelingt es den EvaluatorenInnen, nicht nur Outputs und Outcomes der untersuchten Initiative darzustellen, sondern auch einige Mechanismen zu identifizieren, die für das Handeln der Zielgruppe im Innovationsprozess relevant sind, und damit auch für die genauere Gestaltung des Programms. Hinsichtlich der Programmumsetzung in den betrauten Agenturen werden Prozessveränderungen greifbar, die durch die Initiative ausgelöst wurden bzw. als positive Wirkungen im Sinne der Zielsetzung der untersuchten Initiative gelten können. Es erfolgt eine Programmdarstellung, die die Zielsetzungen des Programms wiedergibt. Sie hätte allerdings noch vertieft werden können, da etwa eine Strukturierung von übergeordneten bzw. längerfristigen und unmittelbar-operativen Zielen nicht vorgenommen wird. Es ergeben sich im Berichtsverlauf immer wieder Unklarheiten, welchen Status bestimmte Daten und Interpretationsweisen für die Analyse der Art und Weise haben, wie das Programm wirkt und mit einzelnen seiner Komponenten zur Erreichung bestimmter Ziele beiträgt. Die Evaluation hätte durch eine eingehendere Auseinandersetzung mit dem von den Verantwortlichen supponierten Wirkmodell bzw. der Rekonstruktion eines sinnvoll annehmbaren Wirkmodells gewinnen können. Es wird im Zuge der Analyse deutlich, dass die Initiative sich in ihrer faktischen Umsetzung ungleichmäßig auf zwei Bezugsprogramme verteilt hat, die in unterschiedlicher Weise operieren. Während die herangezogenen Informationen es ermöglichen, diesen Umstand zu identifizieren und in mehrfacher	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Hinsicht zu beschreiben, wird ihm doch nicht als einem unter Umständen wesentlichen Einflussfaktor auf die Möglichkeiten, intendierte Ziele auch zu erreichen, vertiefend nachgegangen. Für die Analyse des Programmgesehens werden Monitoringdaten und –auswertungen der Agentur direkt übernommen. Auf Schwächen dieser Daten wird von den EvaluatorsInnen hingewiesen. Es werden den Auswertungen, die erhalten wurden bzw. anhand der erhaltenen Daten möglich wurden, jedoch keine neue Auswertungen zur Behebung einiger Schwachpunkte gegenüberstellt. Dies hätte allerdings nur mit beträchtlichem Zusatzaufwand bewerkstelligt werden können, da das gesamte Programmgesehen neu klassifiziert werden hätte müssen. In der Untersuchung der Programmwirkungen wurde eine gut geeignet Methode zur Anwendung gebracht, jedoch hätte offenbar noch stärker in die Erreichung von passenden InterviewpartnerInnen investiert werden können. Die Wirkungsweise des Programms wird anhand der herangezogenen Informationen und der analytischen Vorgehensweise ein Stück weit, jedoch nicht vollständig greifbar. Es wird auf Output-Additionalität Bezug genommen, die sich unter bisherigen Fördernehmern feststellen lässt. Einige intendierte Effekte und potenzielle Faktoren, die auf das Eintreten dieser Effekte Einfluss gehabt haben könnten, können anhand der Daten zu Fördernehmern und Förderwerbern jedoch nur unvollständig bzw. nicht untersucht werden. Zusätzliche Daten auch von Mitgliedern der Zielgruppe, die im Beobachtungszeitraum keinen Projektantrag eingereicht oder gefördert erhalten haben, hätten eine vollständigere und tiefergehende Analyse ermöglichen können. Wissenschaftliche Studien zum Themenbereich der untersuchten Initiative, die mit zu den Outputs dieser Initiative zählen, werden in mehreren Analyseschritten herangezogen, allerdings nicht in allen. Inwiefern eine noch intensivere bzw. über weitere Analyseschwerpunkte erstreckte Bezugnahme möglich gewesen wäre, hängt von den konkreten Inhalten dieser Studien ab und kann nicht beurteilt werden. Darüber hinaus reichende Fachliteratur, die den Operationsbereich des Programms noch näher beleuchten hätte können, um Bedingungen und Einflussfaktoren zu erkennen, die die Erreichung von Zielen beeinflussen können, wurde nicht herangezogen. Aus einer differenzierten Analyse von Daten zu Programmwirkungen wird eine Bandbreite an unterschiedlichen Wirkungsweisen der eingesetzten Förderungen sichtbar. Diese Wirkungsweisen werden freilich nicht weiter ausgeleuchtet, wie es z.B. durch eine ergänzende Literaturanalyse über anderweitig erbrachte empirische Erkenntnisse zum Innovationsbereich erfolgen hätte können. Die Ergebnisse aus den verschiedenen eingesetzten Methode werden mit denen anderer Methoden nicht überall in Beziehung gesetzt, wo dies offenbar grundsätzlich möglich gewesen wäre, um noch weiter erhärtete Analyseergebnisse zu erzielen. Ein ausgewiesenes Programmziel besteht darin, Awareness den Zielgruppen für den unterstützten Innovationstyp zu schaffen, sodass diesen Zielgruppen der Charakter potenzieller Neukunden der betrauten Agentur zukommt, sofern sie sich um eine Förderung im evaluierten Programm bewerben. Es werden allerdings nur NeukundInnen der Agentur untersucht, sodass ,verschiedene mögliche Beiträge zum Auslösen eines bestimmten Typs von Innovationsaktivitäten mit dem Neu-Tätigwerden von Unternehmen gleichgesetzt wird, das für die Agentur durch die Antragstellung auch erkennbar ist. Die Evaluation orientiert sich hier stärker an einer Formaldefinition der Programmformulierung als an Logiken, die in breiterer Weise den Beitrag der Maßnahme zum interessierenden Innovationsgeschehen aufzeigen können.	
N5 Transparenz von Werten	Eine explizite Auseinandersetzung mit den Bewertungsgrundlagen, anhand derer der untersuchten Initiative Wert zugemessen wird, wird den BerichtsadressatInnen nicht angeboten. Die Analyse ist unter anderem an Zielerreichungen orientiert. Zur genaueren Art der Auseinandersetzung werden allerdings keine näheren Erläuterungen gegeben. Es wird eine Analyse von	4

	Zielerreichungen vorgenommen, die einige Ziele herausgreift, die die EvaluatorInnen als die wesentlichsten betrachten. Daneben wird eine Reihe anders gelagerter Betrachtungen angestellt. Eine systematisch-objektive Überprüfung von Zielerreichungen, wie sie etwa mit Messungen innerhalb eines Kontrollgruppen-Designs vorliegen würde, wird nicht verfolgt. In abwägendem Stil werden an verschiedenen Berichtsstellen unterschiedliche Überlegungen präsentiert, die Bezugspunkte für die Einschätzung von Stärken und Schwächen von Konzeptualisierung und Umsetzung der Initiative bieten können. Diesbezüglich ermangeln allerdings Klarheit und leichte Nachvollziehbarkeit. Diese Überlegungen sind über zahlreiche Passagen verteilt und es wird letztlich kein klares Gesamtbild erzeugt, für welche Bewertungsmaßstäbe sich die Evaluation entschieden hat. Manche der dargestellten Argumentationsstränge finden keinen Niederschlag in den zentralen Schlussfolgerungen und Empfehlungen. Damit könnten BerichtsadressatInnen den Bericht auch als eine Art Auswahlmenü wahrnehmen, bei dem man einzelnen Hinweisen folgen kann und anderen nicht. Die zentralen Schlussfolgerungen transportieren die Hinweise, die mit Bewertung zu tun haben, nur in Ausschnitten. In die Analyse fließen Sichtweisen der verschiedenen in die Datenerhebungen einbezogenen Akteursgruppen ein. Es wird allerdings keine Reflexion angeboten, welche Sichtweisen an welcher Stelle zum Tragen kommen sollen, ob die verschiedenen Sichtweisen auch unterschiedliches Licht auf die untersuchte Initiative werfen, und welche Bedeutung dies gegebenenfalls für deren Gesamteinschätzung haben könnte.	
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist in allgemein verständlicher Sprache abgefasst, ohne die AdressatInnen mit fachsprachlichen Termini oder Technismen zu belasten. Tabellen und Grafiken werden zur Vermittlung wesentlicher Informationen eingesetzt und sind anhand der beigegebenen Tabellen- und Abbildungsverzeichnisse leicht auffindbar. Im ersten Berichtsabschnitt werden Programmanlage, zentrale Ergebnisse und Empfehlungen in gut strukturierter Weise vermittelt. Ein Executive Summary ist nicht beigegeben. Das Programm wird hinsichtlich Konzeption, Zielsetzung und Instrumentierung klar und unter Bezug auf Programmunterlagen dargestellt, sodass auch bisher nicht eng mit dem Programm Vertraute der Analyse folgen können. Die Bezugsprogramme, die durch die untersuchte Initiative ergänzt werden, werden ebenfalls dargestellt. Allerdings werden in der Darstellung der verschiedenen Analyseschritten zahlreiche Nebenbemerkungen angebracht, die für BerichtsadressatInnen nur schwer zu verfolgen sind und deren Gehalt sich für sie kaum zu einem Gesamtbild fügen dürfte, insbesondere wenn ihnen nicht viel Zeit für die Lektüre zur Verfügung steht.	2
N8 Nutzung und Nutzen der Evaluation	Aus Angaben zum Methodeneinsatz geht hervor, dass einige VertreterInnen von Stakeholder-Organisationen im Operationsbereich des Programms sowie zahlreiche Mitglieder der Zielgruppen im Zuge von Datenerhebungsverfahren auf die Evaluation aufmerksam gemacht wurden, sodass sie grundsätzlich ein Interesse an deren Ergebnissen entwickeln konnten. Mitglieder der Zielgruppen wurden allerdings nur dann erreicht, wenn sie vom Programm bereits Kenntnis hatten (Fördernehmer und nicht erfolgreiche Antragsteller). Dass eigens ein Kick-Off mit den AuftraggeberInnen erwähnt wird, könnte einen Hinweis darstellen, dass weitere Interaktionen nicht stattgefunden haben. Dies muss allerdings als eine ungesicherte Annahme betrachtet werden und wird in der Einschätzung nicht berücksichtigt. Konkrete Vorstellungen oder Pläne, wie die Evaluationsergebnisse von wem genutzt werden sollen, werden nicht dargestellt.	3
F3 Vollständige und faire Überprüfung	Der Bericht ist in neutraler Sprache abgefasst und benennt Stärken und Schwächen des untersuchten Programms. Er ist in einer Weise formuliert, die teils Stärken oder Schwächen ausdrücklich benennt, teils aber die herangezogenen Fakten	3

	darstellt ohne dazu leicht erkennbar zu machen, wie die EvaluatorInnen diese Fakten einschätzen. In manchen Analyseabschnitten werden Aspekte der Programmumsetzung und Effekte, die eigentlich Gegenstand von datengestützten Analysen gewesen wären, schlicht festgestellt, ohne Daten bzw. Informationsquellen zu benennen. Die Größenordnungen von beobachteten Effekten werden an mehreren Stellen nicht greifbar gemacht. In manchen Analyseabschnitten werden Sichtweisen der InterviewpartnerInnen über mögliche Ursachen direkt und kommentarlos übernommen.	
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Die Positionierung der Initiative im Portfolio der mit der Umsetzung betrauten Agentur wird mithilfe von Förderdaten dieser Agentur beleuchtet. Die Analyse erscheint allerdings auch in manchen Punkten unvollständig, um den Stellenwert des Programms in seinem Kontext gut verstehen zu können. So wird etwa darauf hingewiesen, dass das vom Programm geförderte Geschehen aus anderen Quellen mitfinanziert wurde, und dass ähnliche Projekte auch in anderen Programmen finanziert wurden, ohne allerdings der Frage nachzugehen, welcher Beitrag zur Erzeugung der beobachteten Outputs und Outcomes der untersuchten Initiative unter diesen Umständen zugerechnet werden kann. Dadurch besteht das Risiko, dass die durch das Programm ausgelösten Effekte falsch eingeschätzt werden. Einschränkungen für die gewählte Herangehensweise liegen hinsichtlich der Erkennbarkeit von Wirkungsvoraussetzungen und potenziellen Einflussfaktoren im Bereich der Zielgruppen vor (vgl. N4). Die Programmanlage wird international kontextualisiert, indem einige Vergleichsprogramme in Fallstudien beleuchtet werden. Die hier erbrachten Informationen reichen aus, um einen Vergleich auf der Ebene grundsätzlich verfolgter Strategien anzustellen, jedoch nicht, um die genaueren Wirklogiken der Vergleichsprogramme und die Rolle von Faktoren zu erfassen, die deren Erfolg jeweils fördern oder behindern. Dazu hätte es wesentlich umfangreicherer Daten und Analysen bedurft.	3
G3 Beschreibung von Zwecken und Vorgehen	Ein Eingangskapitel stellt den Methodeneinsatz dar und wird durch weitere, in Anhängen beigegebene Informationen ergänzt. Mit den vorhandenen Angaben wird eine Einschätzung der Aussagekraft und Tragfähigkeit der erbrachten Ergebnisse ermöglicht. Auf Einschränkungen in Datenbeständen, die manche grundsätzlich als angebracht erachteten Analyseschritte verunmöglicht haben, wird von den EvaluatorsInnen hingewiesen. Das Evaluationskonzept und -design wird allerdings nicht kompakt und übersichtlich dargestellt, sondern bleibt im Zuge der Lektüre des Berichts schrittweise zu entdecken. Überlegungen, warum bestimmte Analyseschritte an welcher Stelle gesetzt werden, werden nicht dargestellt. Es werden keine Überlegungen dazu vermittelt, welche Daten aus Sicht der EvaluatorsInnen wie gut geeignet sind, um verschiedene Teilziele und die auf deren Erreichung bezogenen Programmelemente bzw. Umsetzungsschritte zu analysieren (vgl. auch N4). An verschiedenen Berichtsstellen werden Fragestellungen genannt, zu denen nicht klar ist, ob es sich dem Evaluationskonzept ursprünglich zugrunde gelegte Fragestellungen handelt. Eine transparente Beziehung zwischen diesen Fragestellungen und der Programmanlage bzw. Programmzielsetzungen wird nicht hergestellt. Anhand der isolierten Darstellung der Programmziele ist es für BerichtsadressatInnen grundsätzlich möglich, verfolgte Fragestellungen mit ausgewiesenen Zielen zu vergleichen. Jedoch ist wohl davon auszugehen, dass viele BerichtsadressatInnen diese Mühe eher nicht auf sich nehmen werden, und dass auch eine	3

	leichter auffassbare bzw. konzisere Darstellung möglich gewesen wäre. Die BerichtsadressatInnen werden mit einer breiten Palette an Datenlagen und Interpretationen versorgt, ohne dass für sie gut greifbar gemacht würde, ob und inwiefern dabei Tragfähigkeit und Aussagekraft in unterschiedlichem Ausmaß gegeben sind. Eine Gesamtbetrachtung der Methodik, die Vorteile und etwaige Nachteile erläutert, ist nicht beigegeben.	
G8 Begründete Schlussfolgerungen	Schlussfolgerungen und Empfehlungen finden sich in zwei Kapiteln. In einem dieser Kapitel werden Umsetzungsaspekte des Programms behandelt, und im Anschluss Herausforderungen für die betraute Agentur dargestellt. Das zweite dieser Kapitel stellt die zentralen Schlussfolgerungen getrennt von den Datenergebnissen dar, wobei nach Schlussfolgerungen und daraus ausfließenden Empfehlungen strukturiert wird. Ein transparenter Bezug auf zuvor berichtete Fakten ist weitestgehend vorhanden. Es wird jedoch auch einmal eine Faktenlage angesprochen, über die zuvor nicht berichtet wurde. Nicht alle Hinweise, die zu verschiedenen Aspekten der Programmgestaltung und Programmwirkung in der angestellten Analyse erhalten wurden, finden gleichmäßige Berücksichtigung, sodass manche hier in den Vordergrund gestellten Argumente letztlich auch gegenläufig zu an anderer Stelle vorgebrachten Argumenten sein könnten. Es finden sich darüber hinaus schlussfolgernde Passagen in verschiedenen Berichtsabschnitten, die zum Teil in Bezug auf Fakten transparent sind, zum Teil aber auch als vielleicht doch zu vorschnelle Urteile anzusehen sind.	3

Evaluationsbericht Nr. 16		Beobachtungszeitraum: 2007 - 2010	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Die zentralen Beteiligten der Programmverantwortlichen und Programmumsetzenden wurden in einem Kick-off Meeting einbezogen, das der Abstimmung von Erwartungen an die Evaluation galt, sowie in Zwischenbesprechungen und einem Logic Chart –Workshop. Es kann von einer guten Einbeziehung zentraler Programm-beteiligter gesprochen werden. Aus der vorhandenen Darstellung wird allerdings nicht genau nachvollziehbar, welche Akteure an welcher Stelle teilnahmen und ob eventuell bei einzelnen Schritten relevante Akteure übersehen wurden. Betroffene (Programtteilnehmer bzw. Antragsteller) sowie einige Stakeholder wurden in Datenerhebungsverfahren (strukturierte Befragung, Interviews) einbezogen.	2	
N2 Klärung der Evaluationszwecke	Es wird angegeben: Klärung der Effizienz hinsichtlich der Erreichung des Globalziels des Programms. Identifikation von etwaigen Schwachstellen, um zweckmäßige Modifikationen bzw. Ausbaumöglichkeiten vorzuschlagen. Eine Priorisierung bestimmter eingegrenzter Fragestellungen ist nicht zu erkennen.	2	
N4 Auswahl und Umfang der Informationen	Es werden mehrere gut geeignete Methoden eingesetzt, um die Programmschritte, die das konkrete Fördergeschehen ausgelöst haben, und die Begleitmaßnahmen, die neben dem Förderinstrumentarium eingesetzt wurden, analysierbar zu machen. Genauere Evaluationsfragestellungen werden nicht angegeben. Zu einzelnen Analysepunkten werden die Ergebnisse verschiedener Methoden gemeinsam analysiert (Triangulation). Eine Logic Chart wurde gemeinsam mit zentralen Programm-beteiligten erstellt, um die Programmlogik aufzuarbeiten und strukturiert darzustellen. Programmziele werden aufgegriffen und in übergeordnete und intermediäre Ziele differenziert. Outputs und Outcomes werden gegliedert. Die Logic Chart erfasst allerdings nicht alle Programmelemente, die Anhand der vorfindlichen Programmbeschreibung zur Erreichung von Outputs und Outcomes beitragen (z.B. Juryprozess, Kooperationen mit anderen Akteuren). In der typischen Vorgehensweise des Logic Chart-Modells werden auch etwaige außerhalb des Operierens des Programms angesiedelte Voraussetzungen oder potenziell intervenierende Faktoren (Risiken) nicht verzeichnet. Andere Modelle der strukturierten Beschreibung von Programmlogiken hätten hier noch weiter führen können. Die Logic Chart verzeichnet auch Elemente, die im untersuchten Programm nicht vorlagen, sondern von den EvaluatorInnen im Ergebnis ihrer Analyse für eine zukünftige Umgestaltung des Programms empfohlen werden. Dies wird zwar angemerkt, erscheint aber im Sinne einer analytischen Strukturierung des Evaluations-gegenstands, so wie er zum Evaluationszeitpunkt zu untersuchen war, wenig adäquat. Die Analyse der Wirkungsweise des Programms erscheint in mehrerer Hinsicht beeinträchtigt: • Zur Verfolgung der Wirkungsweise des Programms werden Ergebnisse und Empfehlungen einer zuvor erfolgten Zwischen-evaluation des untersuchten Programms aufgegriffen. Auf Veränderungen des Programms während seiner Laufzeit wird hingewiesen. Es wird jedoch in der weiteren Analyse keine Spiegelung des Programmgeschehens oder seiner Effekte an den umgesetzten bzw. nicht umgesetzten Empfehlungen vorgenommen, und es wird nicht verfolgt, ob bestimmten	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Adjustierungen bestimmte Wirkungen zugeschrieben werden können. Auf einen Teil der Programmaktivitäten, der die Vorgehensweise des Programms seit der vorangegangenen Zwischenevaluation ergänzt hat, wird nicht eingegangen. Anhand der vorfindlichen Darstellungen kann nicht entschieden werden, ob hierfür unzureichende Datenerhebungen ausschlaggebend waren, oder die Analyse trotz erhältlicher Daten verkürzt geführt wurde. • In der Umfrage unter ProgrammteilnehmerInnen kann die Befragungsbasis als nachteilig angesehen werden (erfolgreiche EinreicherInnen von zwei Calls unter drei, sämtliche Einreicher/innen von zwei Calls unter drei, beide Kriterien nur für einen Call erfüllt). Zusatzerhebungen hätten für eine bessere Befragungsbasis sorgen können. • Hinsichtlich der Mobilisierungswirkung des Programms wäre es erkenntnisreicher gewesen, die Bekanntheit und Bedarfsgerechtigkeit des Programms in der gesamten Zielgruppe zu erheben, und nicht lediglich unter vom Programm bereits erreichten Mitgliedern dieser Zielgruppe.	
N5 Transparenz von Werten	Der Bericht handelt die Programmumsetzung gemäß dem Evaluationszweck in gleichsam selbstverständlicher Weise ab. Eine Auseinandersetzung mit Bewertungsmaßstäben, die unter Umständen ein anderes Bild ergeben hätten können, erfolgt nicht. Aus der Erhebung unter Zielgruppen-Mitgliedern werden Ergebnisse berichtet, die auf deren Maßstäben dafür beruhen, dem Programm in unterschiedlicher Weise Wert zuzuschreiben, und so zu einem differenzierten Bild führen. Bewertungsmaßstäbe, die außerhalb der in Programmdokumenten vorliegenden Konzeptualisierung liegen, aber von den EvaluatorInnen für die Reflexion der Programmkonzeption als wesentlich erachtet werden, führen zur Empfehlung, das Programm hinsichtlich eines auf dieser Basis identifizierten Bedarfs zu adjustieren. Wenn der Bericht auch insgesamt ein konsistentes Bild erzeugt, so findet doch eine begründende Auseinandersetzung damit, ob andere Bewertungsmaßstäbe unter Umständen ein anderes Bild ergeben hätten können, und warum sie nicht gewählt wurden, nicht statt. Eine Prominenz des Kriteriums der Zufriedenheit von Mitgliedern der Zielgruppe, die durch das Programm bereits erreicht wurde, gegenüber möglichen andersartigen Herangehensweisen daran, ob und inwiefern das Programm unter Umständen an die angezielte FTI-Landschaft noch besser herantreten könnte, ist festzustellen. Stellenweise wird das Kriterium einer klaren Abgrenzung des Programms herangezogen, das sich in dieser Form aus den Programmzielsetzungen, so wie sie berichtet sind, nicht ergibt.	3
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist gut strukturiert und in einer nicht-technischen Sprache gehalten, die ihn für breite Adressatenkreise auffassbar macht. Zentrale Informationen werden in Grafiken und Tabellen aufbereitet. Bereits das Executive Summary bietet eine Programmbeschreibung, die es auch für nicht eng mit dem Programm Vertraute gut ermöglicht, dem Evaluationsbericht zu folgen. Qualitative Ergebnisse werden adäquat und anhand von Zitaten gut nachvollziehbar wiedergegeben. Eine Logic Chart sorgt für eine gute, strukturierte Auffassbarkeit des Programms. Sie ist allerdings mit einer hypothetischen zukünftigen Gestaltungsweise durchsetzt. Eine getrennte Darstellung von tatsächlich vorliegender Programmverfassung und Vorschlag zur Umgestaltung wäre hier vorzuziehen gewesen. Die Programmdarstellung wird unmittelbar mit Ergebnissen aus Analyseschritten und Bewertungen durch die EvaluatorInnen versehen. Während ein guter Überblick erzielt wird, erschwert diese Darstellungsweise die Nachvollziehbarkeit der einzelnen empirischen und logischen Schritte der Analyse.	3
N8 Nutzung und Nutzen der Evaluation	Auf Grund der beschriebenen Einbeziehung von Beteiligten und Betroffenen und insbesondere von mehrfachen Interaktionen zwischen EvaluatoInnen und AuftraggeberInnen im Evaluationsprozess (vgl. N1) kann davon ausgegangen werden, dass die	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Evaluation gute Voraussetzungen dafür geschaffen hat, für ihre HauptadressatInnen nützlich zu sein. Weitere Angaben zu etwaigen Schritten, die Evaluation auf konkrete Nutzungsweisen durch bestimmte Akteure zu beziehen, finden sich nicht. Einschränkungen ergeben aus dem Ausmaß der Erfüllung mehrerer Nützlichkeits-, Fairness- und Genauigkeitsstandards	
F3 Vollständige und faire Überprüfung	Stärken und Schwächen des untersuchten Programms werden identifiziert und dargelegt. An mehreren Stellen wird in der Interpretationen von Ergebnissen auf mögliche Alternativen hingewiesen. Eine Prävalenz des Kriteriums der Zufriedenheit von Mitgliedern der Zielgruppe, die durch das Programm bereits erreicht wurden (vgl. N5), sowie eine Vernachlässigung von Datenerhebungen, die noch nicht erreichte Mitglieder der Zielgruppe einbeziehen (vgl. N4), kann die Identifikation von Stärken und Schwächen beeinflusst haben. Diesbezügliche Überlegungen werden nicht präsentiert. In der Beschreibung von unter den FördernehmerInnen eingetretenen Wirkungen wird auf deren Wahrnehmung und Darstellung vertraut, ohne dass eine Auseinandersetzung präsentiert wird, ob diese Darstellungen eventuell mit einem Bias behaftet sein könnten. An manchen Stellen wird über Fakten in wertender Sprache berichtet oder es werden Bewertungen ausgesprochen, zu denen die Faktenlage nicht transparent wird.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Im Zuge der Programmdarstellung wird benannt, dass das Programm durch zahlreiche regionale Initiativen zur Förderung desselben Innovationsbereichs kontextualisiert ist, und auch - in einem nicht näher bestimmtem Ausmaß - auf regionaler Ebene mit Netzwerkpartnern kooperiert. Diese Kooperationen scheinen auch in der Logic Chart als eine Programmaktivität in der Verfolgung der Programmziele auf. Auf den Einfluss von Aktivitäten der Partner bzw. auf Synergiebildungen, innerhalb derer sich auch eventuelle Schwächen in der Programmumsetzung identifizieren lassen könnten, sodass sie eine Adjustierung motivieren könnten, wird jedoch im Weiteren nur bedingt analytisch eingegangen. Der vorliegende Bericht konzentriert sich auf den unmittelbaren Verantwortungsbereich des Programms. Hinsichtlich der Programmwirkung in dem Segment der FTI-Landschaft, in dem das Programm Verbesserungen auszulösen versucht, wären umfangreichere Datenerhebungen zielführend gewesen. Es wird darauf hingewiesen, dass das Programm Berührungspunkte und potenzielle Überlappungen mit einem anderen Förderangebot der betrauten Agentur aufweist. Dies wird jedoch nicht zum Gegenstand einer genaueren Analyse.	4
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht stellt der angestellten Untersuchung ein klar erkennbares Methodenkapitel voran. Hier werden die eingesetzten Methoden angeführt, und zur durchgeführten Befragung Befragungsbasen und Rücklaufquoten angegeben. InterviewpartnerInnen werden in einem beigegebenen Anhang gelistet. Der Bericht stützt sich allerdings in wesentlichen Teilen auch auf Aussagen von Betroffenen, zu denen er keine gut nachvollziehbaren methodischen Angaben macht. Verschiedene methodische Hinweise werden gegeben. Die Einschätzbarkeit von Programmwirkungen zum Evaluationszeitpunkt wird in adäquater Weise auf vorliegende Anzeichen bezogen, die Anzeichen für längerfristige Auswirkungen darstellen können. An verschiedenen Stellen – wenn auch nicht durchgehend – wird die Tragfähigkeit von Ergebnissen kommentiert.	3

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Der Bezug zwischen Programmbeschreibung, Evaluationszweck und Evaluationsdesign ist im Wesentlichen erkennbar. Evaluationsfragestellungen, die mithilfe der eingesetzten Methoden jeweils beleuchtet werden sollten, werden allerdings nicht gelistet, und der Bericht stellt keine Überlegungen über Vor- und Nachteile des eingesetzten Methodensets zur Verfügung. Die Logic Chart wird offenbar in erster Linie als Präsentationsmittel verstanden und es wird nicht nachvollziehbar, ob und inwiefern sie dazu genutzt wurde, das Evaluationsdesign zu schärfen. Es wird nicht kommentiert, inwiefern und weshalb der Bericht auf verschiedene Aspekte der aus der Logic Chart entnehmbaren angenommenen Wirklogik unterschiedlich intensiv nachgeht. Gründe für die ungleichmäßige Gestaltung der Befragungsbasis für die Erfassung von Erfahrungen und Sichtweisen von Betroffenen (AntragstellerInnen bzw. FördernehmerInnen) werden nicht angegeben. Nachteile für die Tragfähigkeit der gewonnenen Ergebnisse werden nicht so klar dargestellt, dass sie für alle LeserInnen des Berichts jedenfalls erkennbar würden. Konzeptuell diskutabel erschien die Heranziehung eines Indikators, mit dem die Qualität der eingereichten Anträge beleuchtet werden soll, eine Erläuterung wird jedoch nicht angeboten.	
G8 Begründete Schlussfolgerungen	Schlussfolgerungen werden in einem eigenen, von Ergebnisdarstellungen getrennten und leicht auffindbaren Kapitel präsentiert. Kapitel, die Ergebnisdarstellungen dienen, werden teils durch klar erkennbare zusammenfassende Schlussfolgerungen abgeschlossen. Schlussfolgernde Passagen finden sich darüber hinaus auch anderen Stellen des Berichtstexts im direkten Zusammenhang mit Darstellungen von Faktenlagen. Auch die Programmdarstellung wird bereits mit Bewertungen und Empfehlungen für die weitere Vorgehensweise versehen. Einige Schlussfolgerungen, die die Programmkonzeptualisierung und –umsetzung betreffen, und von manchen AdressatInnen unter Umständen für wesentlich gehalten werden könnten, finden sich unmarkiert im Text und sind nur bei einer durchgehenden Lektüre des gesamten Berichts auffindbar. Die abschließenden Schlussfolgerungen und Empfehlungen sind mit einem klaren, transparenten Faktenbezug ausgestattet. Konzentration liegt über weite Strecken auf einer kompakten Zusammenfassung des in der Umsetzung zustande gekommenen Programmgeschehens. Explizite Empfehlungen werden kaum ausgesprochen, lediglich eine spezifische Empfehlung zur Ergänzung des Programms um eine weitere Maßnahme wird klar herausgearbeitet. Es wird aus der Textierung an mehreren Stellen nicht klar, ob ein Handlungsbedarf geortet wird oder die bisherigen Vorgehensweisen als berechtigt und zielführend bestätigt werden.	3

Evaluationsbericht Nr. 17		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ (Zwischenevaluation)		
Evaluationsschwerpunkte	Relevanz, Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In der Anwendung von Datenerhebungsmethoden (wurden Programmverantwortliche, Programmumsetzende, eine in die Programmumsetzung involvierte Akteursgruppe im Zielgruppenbereich, und weitere Stakeholder (Begünstigte) des Programms einbezogen. In allen Fällen beschränkte sich die Einbeziehung auf einen Ausschnitt aus allen vorliegenden Programmlinien. Soweit der präsentationentnommen werden kann, beschränkt sich die Einbeziehung von Programmumsetzenden offenbar auf einen Ausschnitt aus allen vorliegenden Programmlinien. Weitere Angaben finden sich nicht. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in die Planung der Evaluation hinsichtlich von Informationsbedürfnissen einbezogen worden wären	3	
N2 Klärung der Evaluationszwecke	Angegeben wird: Analyse und Reflexion des bisherigen Programmverlaufs durch die Untersuchung von Konzeption, Umsetzung und bisher feststellbaren Wirkungen, um Schlussfolgerungen und Empfehlungen für die Weiterentwicklung des Programms zu erarbeiten. Dabei wird eine Schwerpunktsetzung auf ausgewählte Programmlinien vorgenommen. Es wird aus der Berichterstattung nicht klar, ob der Evaluationsgegenstand auf die „Schwerpunkte“ eingegrenzt wurde, um die Evaluation des sehr umfangreichen und komplexen Programms handhabbar zu machen, oder ob für diese Schwerpunkte eine intensivere Beleuchtung vorgesehen war, ohne auf ein Gesamtbild zu verzichten.	3	
N4 Auswahl und Umfang der Informationen	Evaluationsfragestellungen, die die Vorgehensweise der Evaluation näher strukturieren, werden nicht angegeben. Die Evaluation stützt sich auf vier parallel eingesetzte Methoden, die grundsätzlich für die Analyse des Programms unter Bezug auf Evaluationszweck und Evaluationsschwerpunkte gut geeignet sind. Evaluationsergebnisse von im untersuchten Programm zusammengeführten Vorgängerprogrammen wurden herangezogen. Die herangezogenen Informationen erlauben in erster Linie eine eingehende Darstellung des Projektgeschehens, das innerhalb der einzelnen Programmlinien beobachtet werden konnte. Die Analyse wird auf zwei der drei Interventionsfelder des Programms beschränkt. Dies kann als sinnvolle Eingrenzung des Evaluationsgegenstands verstanden werden. Allerdings hat sich damit die Evaluation von vornherein der Möglichkeit begeben, das gemeinsame Wirken aller unter dem Programm agierenden Programmlinien mit seinen Synergien bzw. etwaigen Synergieverlusten in umfassender Weise zu verstehen und auf dieser Ebene Faktoren zu identifizieren, die für die Erreichung der übergeordneten Zielsetzung des Programms wichtig sein könnten. Das untersuchte Programm ist von hoher Komplexität gekennzeichnet, nicht zuletzt da es mehrere Vorgängerprogramme zusammenfasst. Eine Grafik veranschaulicht die Verortung der einzelnen Programmlinien in der Gesamt-Herangehensweise des Programms an eine Querschnittsmaterie. Eine Logic Chart oder andere genauere Aufarbeitung der Wirklogik des Programms wurde allerdings nicht erstellt. Das Evaluationsdesign war in mehrerer Hinsicht nicht gut geeignet, im umfassender Weise Hindernisse zu erkennen, die der Erreichung von Programmzielen entgegenstehen und in der Folge Adjustierungen motivieren könnten:	4	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	• Indem sich die Analyse an der Gliederung des Programms in mehrere Programmlinien orientiert, erscheint das Programm lediglich als Aggregat der verschiedenen einzelnen Programmlinien. Es wird dargestellt, dass das Programm Interventionsfelder hat, in denen jeweils mehrere Programmlinien enthalten sind. Ob und wie verschiedene Programmlinien ineinandergreifen, oder welche Rolle sie bei der Erreichung von intendierten Zielsetzungen in einem Gesamtrahmen einnehmen, wird jedoch nicht verfolgt. • Für einzelne Programmlinien wird in der Textierung erkennbar, dass sie einer zweistufigen Logik von direkt Begünstigten und indirekt Begünstigten folgen. Derartigen Logiken und Faktoren, die in einer solchen Wirklogik die Erreichung von Effekten beeinflusst haben könnten, wird nicht nachgegangen. • Wirkungen auf und Einschätzungen durch direkt Begünstigte (Antragsberechtigte bzw. Fördernehmer) und indirekt Begünstigte, auf die sich das Programm gemäß seiner Zielsetzung eigentlich richtet, werden nicht gemäß ihrem unterschiedlichen Stellenwert für die Art und Weise, wie das Programm Wirkungen zu erreichen versucht, differenziert. In dieser Hinsicht erscheint die vorgelegte Analyse dem Programm nicht gerecht zu werden. • Eine analytisch-konzeptuelle Schwäche besteht in der mangelnden Gliederung von Outputs, Outcomes und weiteren Wirkungen, die im gegenseitigen Voraussetzungscharakter begriffen werden und sodann auch Einflussfaktoren zuordnen lassen. Es kann eingeschätzt werden, dass eine genauere Einlassung auf die Wirklogiken der einzelnen Programmlinien jedenfalls hilfreich gewesen wäre, um die Analyse und die Ergebnisaufbereitung besser zu strukturieren. • Daten darüber, in welchem Ausmaß einzelne Zielgruppen von den gesetzten Maßnahmen erfasst wurden, wurden nicht herangezogen. • An manchen Textstellen scheint durch, dass Teile des Programms gemeinsam mit KooperationspartnerInnen umgesetzt wurden. Der Rolle solcher Umsetzungsmodalitäten bzw. deren Stärken und Schwächen, die die Erreichung von Programmergebnissen mit beeinflusst haben können, wird nicht untersucht. Insgesamt vorliegende Probleme in der Bewältigung der Komplexität des Evaluationsgegenstands auf der Ebene des Evaluationskonzepts und seiner weiteren Umsetzung lassen den Schluss zu, dass die Evaluation zu einer besseren Bewältigung der Komplexität des untersuchten Programms konzeptiver Schärfungen und wohl auch höherer Ressourcen bedurft hätte.	
N5 Transparenz von Werten	Es erfolgt keine Darstellung zu Überlegungen und Entscheidungen, anhand welcher Maßstäbe dem Programm Wert zugewiesen werden soll. Die Evaluation geht wie selbstverständlich vor, indem sie die in den gewählten Methoden gewonnenen Daten ausweist. In der Berichtstextierung zeigt sich eine gewissen Variabilität, mit welchen Argumenten Ergebnisse als Stärken und Schwächen eingestuft werden. Dabei kommen unter anderem Kriterien wie ein Alleinstellungsmerkmal des Programms zum Tragen, was allerdings nicht als geeignet erachtet werden kann, um Rückschlüsse auf die tatsächliche Treffsicherheit von Programmkonzept und -umsetzung unter empirisch-konkreten Bedingungen der Realwelt zu erbringen, um dem Programm auf dieser Basis Wert zuzuweisen.	5
N6 Vollständigkeit und Klarheit der Berichterstattung	Es liegt ein klar strukturierter Bericht vor, in dem sich AdressatInnen leicht zurecht finden können. Wesentliche Inhalte werden in Grafiken und Tabellen dargestellt. Dem Bericht ist eine achtseitige Zusammenfassung vorangestellt, die den Charakter eines Kurzberichts hat. Die Darstellung der Ergebnisse zu den einzelnen Programmlinien ist übersichtlich und gut auffassbar. Ergebnisse aus unterschiedlichen eingesetzten Methoden werden hier zusammengezogen. In den überblicksartigen Darstellungen werden jedoch teils keine präzisen Angaben über Verteilungen gemacht, sodass für genaue Information doch in den Bericht hineingegriffen werden muss.	2

	Das Programm wird in seiner grundsätzlichen Anlage mit Interventionsfeldern und Programm-linien bereits im Executive Summary überblicksartig dargestellt, sodass ein guter Einstieg in die Analyse auch für nicht eng mit dem Programm Vertraute ermöglicht wird. Im Bericht selbst finden sich kompakte Darstellungen der einzelnen Programmlinien mit zugehörigen Ergebnissen auf Output- und Outcome-Ebene. In den Darstellungen zur Programmumsetzung finden sich allerdings teils Passagen, die Insider-Wissen über die Maßnahmen bzw. deren Vorgänger voraussetzen.	
N8 Nutzung und Nutzen der Evaluation	Durch die Einbeziehung von Programmverantwortlichen, Programmumsetzenden und Betroffenen im Rahmen von Datenerhebungsverfahren (vgl. N1) wurden Voraussetzungen dafür geschaffen, dass diese sich für die Evaluation und ihre Ergebnisse interessieren. Weitere Angaben zu etwaigen Schritten, die Evaluation auf konkrete Nutzungsweisen zu beziehen und diese zu unterstützen, finden sich nicht. Schwächen bei der Umsetzung mehrerer Standards schränken die Nützlichkeit ein. Eine nicht hinsichtlich aller Programmbestandteile gleich systematische Analyse sowie Schwächen in der Übersetzung der angegebenen Zielsetzung der Evaluation und der Evaluationsschwerpunkte in ein überzeugendes und stringent umgesetztes Analysekonzept sind dabei als wesentlich anzusehen.	3
F3 Vollständige und faire Überprüfung	Stärken und Schwächen des untersuchten Programms werden identifiziert. In der Berichterstattung werden Ergebnisse im Tonfall einer neutralen Dokumentation dargestellt. Allerdings werden dabei teilweise Ergebnisse, die auf vorliegende Schwächen in der Programmumsetzung und -konzeption hindeuten könnten, nicht immer auch klar benannt. Auch werden Ergebnisse, die auf Schwächen hindeuten können, selten durch Grafiken unterstützt. Dies kann die Auffassbarkeit gerade für weniger mit dem Programm vertraute BerichtsadressatInnen und eilige LeserInnen beeinträchtigen. Darstellungen zur Programmumsetzung und Outputs erfolgen teilweise in einer unmittelbar wertenden Formulierungsweise. Während sich der Bericht stark auf Sichtweisen von Programmbeteiligten und Betroffenen (Zielgruppen) stützt, ist an mehreren Stellen nicht evident, ob es sich bei Aussagen des Berichts um Einschätzungen von Beteiligten und Betroffenen oder Aussagen der EvaluatorenInnen auf Basis ihrer eigenen Sichtweisen handelt.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Die Untersuchung geht nicht auf die Umwelten ein, in denen das untersuchte Programm seine Wirkungen zu erzeugen versucht, und verfolgt auch nicht die wechselseitige Kontextualisierung mehrerer Programmlinien in Interventionsfeldern des untersuchten Programms.	5
G3 Beschreibung von Zwecken und Vorgehen	Ein eigenes Methodenkapitel steht am Anfang des Berichts. Grundsätzliche Überlegungen zur möglichen Verfolgung von Wirkungen und zur Evaluierbarkeit zum erreichten Entwicklungsstadium des Programms werden dargestellt, und es wird zurecht eine nur eingeschränkte Analysierbarkeit zum gegebenen Evaluationszeitpunkt festgestellt. Da Evaluationsfragestellungen nicht angegeben werden, ist dennoch nicht immer gut nachvollziehbar, welche Methoden mit ihren Ergebnissen in Bezug auf welche Evaluationsschwerpunkte Erkenntnisse bringen sollen. Das Programm wird einer nur ausschnittshaften und ungleichmäßig intensiven Untersuchung unterzogen, was zwar entnommen werden kann, aber nicht durchgehend ausreichend klar herausgestellt wird. Es werden keine Hinweise angeboten, welche	3

	Programmaspekte mithilfe welcher Methoden wie gut analysiert werden konnten und daher besser oder weniger gut eingeschätzt werden können. Wesentliche Teile der herangezogenen Information entstammen den verfügbaren Monitoringdaten. Es findet sich keine leicht auffassbare Auseinandersetzung mit deren Qualität und Tragfähigkeit. Es kann davon ausgegangen werden, dass die AuftraggeberInnen des Berichts über die Fachkenntnisse verfügen, die eine irrtumsfreie Auffassung ermöglichen. Für BerichtsadressatInnen, die nicht bereits mit der Evaluation und ihrem Gegenstand eng vertraut sind, hätte es hier aber klarerer Hinweise und Explikationen bedurft.	
G8 Begründete Schlussfolgerungen	Die einzelnen Abschnitte zu den verschiedenen Programmlinien münden jeweils in eine Aufbereitung von sichtbar gewordenen Stärken und Schwächen, die Schlussfolgerungs-Charakter haben. Dies ist auf Grund der systematischen und gleichmäßig durchgehaltenen Berichtsgestaltung für die AdressatInnen erkennbar und gut zu handhaben, da die Schlussfolgerungen im direkten Konnex mit den zugrunde gelegten Faktenlagen dargestellt sind. Nicht alle Schlussfolgerungen sind allerdings hinsichtlich zugrunde gelegter Fakten transparent und hinsichtlich der Adäquanz der vorgetragenen Argumentation gut nachvollziehbar. Im abschließenden Schlussfolgerungs-Kapitel werden im Zuge von Hinweisen für die weitere Programmgestaltung teils auch neue Fakten angezogen, die in den vorangegangenen Ergebniskapiteln so noch nicht dargestellt sind. Dadurch wird teilweise auch ein anderes Licht auf die Maßnahme geworfen. Wenn auch zweifellos zuzugestehen ist, dass nicht alle Aspekte, die in Detaildarstellungen der Dateninterpretation Erwähnung finden, in die zentralen Schlussfolgerungen aufgenommen werden können, so kommt es doch streckenweise zu Ungewichtungen, was die Einschätzung bestimmter Umsetzungsaspekte betrifft.	4

Evaluationsbericht Nr. 18		Beobachtungszeitraum: 2007 – 2010	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ		
Evaluationsschwerpunkte	Relevanz, Effektivität, Effizienz, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	In Datenerhebungsverfahren wurden Programmverantwortliche und –umsetzende sowie etliche Stakeholder einbezogen. Die beigegebene Liste der InterviewpartnerInnen zeigt, dass einige wesentliche Akteure aus dem FTI-Governancesystem sowie einige wenige Mitglieder der Zielgruppen auf diesem Weg einbezogen waren..Betroffene (Antragsteller und Fördernehmer) wurden in einer Umfrage einbezogen, wobei sich die Abdeckung für zwei Programmteile unterschiedlich darstellt. Angaben zur Interaktion mit dem Auftraggeber, die über das Stattfinden eines Kick Off-Meetings hinausreichen, oder Angaben zur Art und Weise wie Stakeholder identifiziert und ausgewählt wurden, werden nicht gemacht.	2	
N2 Klärung der Evaluationszwecke	Es wird angegeben: Konzeption, Umsetzung, Prozesse und Organisation, Zielerreichung und Wirkungen des Programms, soweit diese zum Evaluationszeitpunkt feststellbar sind, untersuchen. Es werden zwei Hauptfragestellungen genannt, die allerdings nicht in die formulierten Untersuchungsdimensionen eingeordnet werden. Hieraus soll erschlossen werden, ob und wie das Steuerungsinstrument zukünftig fortgesetzt werden soll. Im ersten Berichtsabschnitt findet sich sodann eine Darstellung der Vorgehensweise der Evaluation, die in der Benennung der Untersuchungsschwerpunkte mit der oben genannten nicht vollkommen konsistent ist. Die Durchführung der Evaluation wird formal auf den Evaluationsauftrag der Programmrichtlinien bezogen. Die angegebenen Untersuchungsdimensionen sind äußerst umfangreich, und es findet keine Priorisierung statt, um die zu führende Analyse mit mehr Richtung und Klarheit auszustatten. Die sehr breit angelegte Untersuchung erscheint mit der in der Angabe ihres Vorhabens anzutreffenden Kombination von Evaluationsschwerpunkten, Analysebereichen und dem Kriterium der Zielerreichung unscharf formuliert.	4	
N4 Auswahl und Umfang der Informationen	Die Untersuchung nimmt von einer Programmbeschreibung und einer Darstellung der Programmziele ihren Ausgang. Eine strukturierte Gliederung in übergeordnete, intermediäre und operative Ziele findet dabei allerdings nicht statt. Es werden mehrere Methoden eingesetzt, die für die Durchführung der geplanten Analyse gut geeignet erscheinen. Im Verlauf der Analyse werden zu den angegebenen Untersuchungsschwerpunkten (vgl. N2) jeweils einige Gesichtspunkte verfolgt, die jedenfalls als sachdienlich zu erachten sind. Es bleibt allerdings im Sinn einer übergreifenden Systematik des Vorgehens unklar, ob es sich bei diesen Analysesträngen um die jeweils bestgeeigneten Auseinandersetzungen mit den Outputs, Outcomes und Prozessaspekten der Programmumsetzung handelt. Irritationen entstehen unter anderem aus folgenden Berichtsmerkmalen: • Mit der Analyse von neuen Fördernehmern bei den umsetzenden Agenturen verfolgt die Analyse ein in den Programmzielen niedergelegtes Kriterium, doch werden andere, aus den Programmzielen mindestens ebenso berechtigten bzw. gebotenen Analyserichtungen nicht in derselben Intensität verfolgt. Damit könnte letztlich nicht gut erkannt worden sein, ob das Programm neuartiges FTI-Geschehen ausgelöst hat, und inwiefern die Beobachtungen zum Fördergeschehen dem Programm kausal zurechenbar sind.	3	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	• Da nicht untersucht wird, ob die bei einer Agentur auftretenden Fördernehmer auch bei der jeweils anderen Agentur auftreten, kommt es zu keinem vollständigen Gesamtbild des Fördergeschehens über beide Agenturen hinweg. • Auch in anderen Auswertungsschritten zu den Monitoringdaten des Programms kommt es nicht in allen Punkten zu einer vollständigen Gesamtsicht. Die EvaluatorsInnen halten selbst fest, dass dazu spezifische Auswertungen nötig gewesen wären, die jedoch nicht durchgeführt wurden. In der gewählten Analysestrategie wird so an relevanten Punkten nicht erkennbar, inwiefern die verschiedenen Programmschienen in ihrer Unterschiedlichkeit (Schwerpunktsetzungen, Antragserfordernisse, Fördermodi, etc.) zu einem Gesamt-Programmerfolg beigetragen haben, und inwiefern hier eventuell Adjustierungen ratsam werden hätten könnten. • Umgekehrt wird in der Interpretation von Zufriedenheiten mit der Programmabwicklung nicht nach den umsetzenden Agenturen und nach unterschiedlichen Segmenten der Zielgruppe differenziert. Dadurch können mögliche Einsichten, die für die weitere Gestaltung des Programms wesentlich sein könnten, nicht gewonnen worden sein. • Die genaue Konstitution der Befragungsbasis (Antragsteller und Fördernehmer für eine Agentur, für die andere Agenturen nur Fördernehmer) scheint bei der Ergebnisinterpretation zu manchen Fragestellungen nicht ausreichend berücksichtigt zu werden, sodass die Gefahr verzerrter Ergebnisse besteht. Es kann aus der vorliegenden Präsentation heraus nicht entschieden werden, ob es sich nur um eine Darstellungsschwäche handelt und also die Daten doch adäquat gehandhabt wurden. • Auf wesentliche Umsetzungsaspekte wird eingegangen, doch wurden nicht alle Elemente, die der Bericht selbst herausarbeitet, auch konsequent verfolgt (z.B. Rollen von Programmbeirat und anderen Gremien in Entscheidungen, die auf das Programmgeschehen Einfluss hatten). Auf einige virulente Zonen der Programmumsetzung wird hingewiesen, doch ist letztlich wohl nicht davon auszugehen, dass die Programmumsetzung vollständig untersucht wurde. • Zur fundierten Einschätzung der Rolle der Förderungen für die Zielgruppen im Verhältnis zu anderen existierenden Förderungen hätte ein Vergleich mit Angehörigen der Zielgruppe, die keine Förderung aus dem untersuchten Programm beantragt haben, umfassendere Einsichten ermöglicht. Der Bericht unterhält keine Bezug auf wissenschaftliche Literatur, die zum Themenbereich des untersuchten Programms grundsätzlich vorliegt und dazu beitragen hätte können, die Ergebnisinterpretation zu unterstützen und zu vertiefen. Eine Logic Chart oder andere Darstellungsform der Programmlogik, die den Stellenwert einzelner Beobachtungen aufzeigen hätte können, wurde nicht erstellt. Eine eingesetzten Methoden ist ambitioniert, aber eindeutig im Rahmen der verschiedenen Untersuchungsdimensionen und Evaluationsschwerpunkte von vornherein zu einer unzureichenden Durchführung verurteilt, um für eine analytische Auseinandersetzung mit dem untersuchten Programm wirklich fruchtbar werden zu können. Die EvaluatorsInnen statten dementsprechend die Interpretation mit einer durchaus angebrachten Kautel aus. Es hätte hier einer wesentlich breiteren Informationsbasis bedurft, um mehr zu erzielen als die erbrachten spärlichen Ergebnisse.	
N5 Transparenz von Werten	Es werden keine expliziten Überlegungen zu den Maßstäben dargestellt, anhand derer dem Programm Wert zugewiesen wird. Die verschiedenen Datenquellen werden wie selbstverständlich herangezogen, ohne zu erläutern, was warum aus welchen Ergebnissen geschlossen werden soll. Bei der Überprüfung der Erreichung von Programmzielen finden keine Techniken wie 'Thresholding' Anwendung. Es wird nicht	4

	nachvollziehbar, dass bei Einschätzungen von Datenergebnissen von klar verankerten Kriterien ausgegangen worden wäre, und es werden z.B. Werte in der Bandbreite von ca. 65% bis ca. 75% als Zielerreichungen interpretiert, ohne eine Explikation beizustellen. In verschiedenen Berichtsabschnitten kommen unterschiedliche Maßstäbe zur Anwendung (NeukundInnen der Agenturen, Kundenzufriedenheiten mit der Betreuung durch die Agenturen, etc.), ohne zu erläutern, warum dies die geeignetsten Maßstäbe sind, und wie sie sich zu einer Gesamteinschätzung fügen. An manchen Stellen wird die Perspektive der umsetzenden Agenturen auf das FTI-Geschehen in den Vordergrund gestellt. Das FTI-Geschehen als solches und die Wirkungen, die das Programm darauf erzielt, steht nicht immer im Mittelpunkt, ohne dass diese Perspektivenwahl erläutert wird.	
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist klar strukturiert und stellt Ergebnisse in zahlreichen gut gestalteten Tabellen und Grafiken dar, die anhand der Verzeichnisse leicht auffindbar sind. Technische Materialien werden getrennt in Anhängen beigegeben. Ein ausführliches Executive Summary, das eher einen Kurzbericht darstellt, vermittelt die Programmanlage und Eckdaten des Programmgeschehens, die Vorgehensweise der Evaluation, zentrale Ergebnisse und Empfehlungen. Die berichteten Daten sind mit klaren Quellenangaben versehen. Andere wesentliche Teile des Berichts, wie etwa die Programmdarstellung oder Übernahmen von Ergebnissen anderer Evaluationen, kommen allerdings ohne Benennung von Belegen aus und sind daher weniger transparent. Das Programm wird mit seiner Motivation so dargestellt, dass auch nicht mit ihm Vertraute dem Evaluationsbericht folgen können. Bei Ergebnisinterpretationen wird das Öfteren nicht klar ausgedrückt, was berichtete Daten über das untersuchte Programm letztlich aus Sicht der EvaluatorenInnen aussagen. Der Bericht hält sich häufig in Hinweisen auf, bei denen es den AdressatInnen überlassen bleibt, welche Bedeutung sie ihnen beimessen und welche Schlüsse sie daraus ziehen. Im Berichtstext werden wiederholt Ergebnisse mit verschiedenen Überlegungen verbunden, die in ihrem Vortrag den Status der Ergebnisse in einem analytischen und bewertenden Konzept (vgl. N4, N5) und ihre Lesart durch die EvaluatorenInnen nicht leicht nachvollziehbar machen bzw. zu keinem klaren, den AdressatInnen unmissverständlich angebotenen Resultat führen. Die Schlussfolgerungen sind an Zielformulierungen des Programms ausgerichtet und liefern Erkenntnisse, die zu einer Beurteilung des Ausmaßes der Zielerreichung dienen können, sprechen aber dazu selbst keine Einschätzungen aus.	2
N8 Nutzung und Nutzen der Evaluation	Programmverantwortliche und –umsetzende sowie etliche Akteure des FTI-Governancesystems wurden im Rahmen von Datenerhebungen einbezogen, ebenso wie Begünstigte des Programms. Während für erstere Gruppe damit wohl die Grundlage geschaffen wurde, sie für die Evaluation und ihre Ergebnisse zu interessieren, so kann das für die zweite Gruppe als weniger wahrscheinlich erachtet werden. Hinweise auf weitere Schritte, AdressatInnen für Ergebnisse zu interessieren, finden sich nicht. Es kann davon ausgegangen werden, dass der Bericht mit seiner Vielzahl an Informationen und Überlegungen die HauptadressatInnen dabei unterstützen kann, sinnvolle Veränderungen vorzunehmen und neue Sichtweisen auf das Programm zu entwickeln. Einschränkungen für bestmögliche Nützlichkeit ergeben sich jedoch aus mehreren Standards.	3
F3 Vollständige und faire Überprüfung	Der Bericht pflegt eine neutrale Sprache. Er weist sowohl auf Stärken als auch auf Schwächen hin, die allerdings kaum explizit so genannt werden, sondern eher durch genaues Lesen zu erkennen bleiben. Zu den Programmzielen werden Ergebnissen nicht in einer übersichtlichen und klaren Form korreliert. AdressatInnen des Berichts müssen so aus verschiedenen Berichtsteilen Aussagen	3

	zusammentragen und mit den Zielformulierungen vergleichen, um festzustellen, welche Erkenntnisse jeweils vorliegen. Es ist nicht an allen Stellen klar, wie die Daten interpretiert werden, und warum sie in der vorliegenden Weise interpretiert werden. Daten, die lediglich in Grafiken enthalten sind, aber nicht textlich aufgegriffen werden, könnten unter Umständen Hinweise enthalten, die für das Operieren und die Zielerreichung des Programms wesentlich sind, und es unter Umständen auch in einem anderen Licht erscheinen lassen, als es in der vorfindlichen textlichen Interpretation der Fall ist. In manchen Berichtsabschnitten teilen die EvaluatorInnen nicht mit, ob sie berichtete Erkenntnisse für Hinweise auf Stärken oder auf Schwächen des Programms halten. Auch im Executive Summary wird bei mehreren Ergebnisdarstellungen nicht klar ausgesprochen, ob die EvaluatorInnen sie in ursächlicher Verbindung mit Schritten der Programmumsetzung sehen, oder eher als Verhaltensweisen der Zielgruppen, die dem Programm nicht zurechenbar sind. Insgesamt ist eine gute Auffassbarkeit davon, zu welchen Schlüssen die EvaluatorInnen tatsächlich kommen, an verschiedenen Stellen beeinträchtigt. Es kann davon ausgegangen werden, dass dies die Nützlichkeit des Berichts vor allem für Personenkreise beeinträchtigt, die sich jenseits der unmittelbaren Auftraggeber-Personen ansiedeln und nicht mit den EvaluatorInnen im Austausch stehen.	
F5 Offenlegung der Ergebnisse	Der Bericht enthält auch ein Executive Summary in englischer Sprache, sodass die Evaluation und ihre Ergebnisse international aufgegriffen werden können. Dass in den Empfehlungen die Nummerierung unstimmt ist, könnte unter Umständen einen Hinweis darstellen, dass der Bericht nicht in vollständiger Fassung publiziert wurde. Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Die Evaluation bezieht sich explizit auf den Umstand, dass das untersuchte Programm durch andere Programme innerhalb und außerhalb der umsetzenden Agenturen kontextualisiert ist. Indem es an zwei bereits existierenden Programmen ansetzt, erscheint eine darauf bezogene Analyse dringend geboten und verbindet sich unmittelbar mit der Hauptfragestellung nach der Relevanz des Programms. Die Evaluation wird dem in ihrer Herangehensweise gerecht. Ein eigenes Kapitel widmet sich der Positionierung und Einbettung des Programms in der österreichischen Programmlandschaft. Hier kommt es zu Überlegungen über Reperkussionen auf die Positionierung anderer Programme der umsetzenden Agenturen. Die für die Analyse der Positionierung des Programms herangezogenen Förderungsdaten sind auch geeignet, in einer Portfolio-Betrachtung interessierende Förderkarrieren von Zielgruppen-Segmenten und kumulative Förderungen sichtbar zu machen. Dies steht in direktem Bezug zu einem dargestellten Programmziel. Die Analyse wird allerdings nicht auf eine Ebene geführt, auf der eingeschätzt werden könnte, inwiefern die kumulativen oder synergetischen Gesamt-Förderungen nicht Mitnahmeeffekte darstellen, sondern doch im Sinne von anderen, ebenso existierenden Zielsetzungen des untersuchten Programmes sein könnten. Dazu hätte allerdings das Projektgeschehen aller Kontext-Programme analysiert werden müssen. Hinsichtlich von Klärungen, wie sich das Programm tatsächlich empirisch beobachtbar positioniert, werden zum Teil Daten herangezogen, die lediglich die Programmanlagen und Antragserfordernisse widerspiegeln. Insgesamt wird nicht klar strukturiert bzw. reflektiert, welche beobachteten Eigenschaften des Programms in seinen realen Umwelten in welcher Weise zu seiner besseren oder schlechteren - und in diesem Sinne verbesserungsfähigen - Entwicklung beitragen. (z.B. wird zwar über die Anforderungshöhe der Antragstellung	2

	berichtet, es werden aber keine Überlegungen dazu angeboten, inwiefern eine anforderungsreichere oder niederschwelligere Verfahrensgestaltung zur Erreichung der Programmziele beitragen können).	
G3 Beschreibung von Zwecken und Vorgehen	Ein eigenes Methodenkapitel beschreibt die Einsatzpunkte der meisten gewählten Methoden im Gesamtrahmen der Untersuchung. Eine Grafik erzeugt zusätzliche Übersicht über das Methodenset. Ein Bezug zwischen Berichtsabschnitten und den Methoden, auf deren Ergebnisse sie sich vorrangig stützen, wird hergestellt. Eine klare Programmdarstellung bildet den Ausgangspunkt der Untersuchung. Zur durchgeführten Befragung werden Eckdaten zum erreichten Rücklauf dargestellt, die eine Einschätzung der Tragfähigkeit der erzielten Ergebnisse erlauben. Detaillierte Hinweise, die die Aussagekraft von präsentierten Daten betreffen, werden teilweise auch in Fußnoten oder im Interpretationstext gegeben. Die eingesetzten Fragebögen sind in einem Anhang beigegeben. Reflexionen über etwaige Schwächen der erhaltenen Monitoringdaten für die Einlösung analytischer Fragestellungen werden allerdings nicht präsentiert. In manchen Passagen werden Daten referiert und verschiedene Überlegungen dazu herangezogen, welcher Stellenwert in der Wirkungsweise des Programms ihnen zugemessen werden kann. Eine übergreifende konzeptuelle Explikation ist jedoch nicht vorhanden. Hinichtlich des verfolgten Analysekonzepts ist trotz aller vorhandener Darstellungen doch ein klarer Konnex von Programmkonfiguration, Evaluationszwecken, Evaluationsfragestellungen und jeweils zu deren Beantwortung herangezogenen Methoden nicht vollkommen gegeben. Stellenweise wird erwähnt, dass auf Evaluationsfragestellungen der Ausschreibung eingegangen wird, doch werden diese nicht explizit dargestellt. Insgesamt ist nicht klar, ob und wie Evaluationsfragen strukturiert und gewichtet wurden, und so zur vorliegenden Analyse geführt haben. Dies beeinträchtigt die Nachvollziehbarkeit der Vorgehensweise der Evaluation für alle AdressatInnen jenseits der direkten AuftraggeberInnen. Der Zugang zur Analyse von Effizienz wird nicht hinsichtlich etwaiger Nachteile kommentiert. Es wird nicht berichtet, ob und inwiefern Einschränkungen der Aussagekraft von in der Methodenanwendung tatsächlich erhaltenen Daten vorliegen, und welche Auswirkungen das auf die Analysierbarkeit der verschiedenen Einsatzpunkte der Evaluation gehabt haben könnte.	3
G8 Begründete Schlussfolgerungen	Der Bericht trennt die Ergebnisdarstellungen klar ersichtlich von einem Kapitel für Schlussfolgerungen und einem weiteren Kapitel mit Empfehlungen. Die Schlussfolgerungen sind großteils mit unmittelbaren Verweisen auf die ihnen zugrunde liegenden Fakten ausgestattet. Es werden jedoch auch mehrmals Fakten angezogen, die zuvor nicht berichtet wurden. Es werden auch potenzielle Konsequenzen bestimmter Richtungsentscheidungen dargestellt. Es entsteht jedoch nicht der Eindruck, dass zu allen im Gesamtbericht untersuchten Aspekten in den Schlussfolgerungen und Empfehlungen Handlungsalternativen diskutiert werden. Zu einigen Programmzielen findet sich keine direkte Aussage. Manche Schlussfolgerungen beruhen doch auf freien Überlegungen der EvaluatorInnen, oder kommen im Argumentationsgang ohne Bezugnahme auf an dieser Stelle ebenfalls in Betracht kommende Ergebnisse der gesamten zuvor berichteten Analyse aus	2

Evaluationsbericht Nr. 19		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	formativ-summativ		
Evaluationsschwerpunkte	Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen von Datenerhebungsverfahren wurden Begünstigte des Programms, sowie Mitglieder der Zielgruppe, die am Programm nicht teilgenommen haben, und in großem Umfang einbezogen. Die Identifikation dieser Personen erfolgte durch Auswertung objektiver Daten. Weitere Stakeholder des Programms wurden in Form eines Workshops einbezogen, wobei genauere Angaben zu Umfang und Zusammensetzung dieser Gruppe allerdings nicht gemacht werden. Angaben zur Art und Weise bzw. Intensität, mit der die Programmverantwortlichen einbezogen wurden, werden nicht gemacht. Es findet sich kein gesicherter Hinweis, dass FTI-politische Akteure einbezogen worden wären, die nicht unmittelbar mit dem Programm befasst waren, aber an der Evaluation und ihren Ergebnissen ein Interesse haben könnten.	2	
N2 Klärung der Evaluationszwecke	Angegeben wird: Bereitstellung von Information für die Weiterentwicklung und Weiterführung des Programms für den Programmeigentümer. Bereitstellung von Politikinformation für AdressatInnen auf nationaler und europäischer Ebene. Die Evaluationszwecke sind somit nur in recht allgemein-abstrakter Form umrissen. Eine Adressierung zweier politisch-administrativer Sphären wird festgehalten, wobei allerdings nicht näher hinsichtlich von Zuständigkeiten oder Hauptverantwortlichen (z.B. Gremien) spezifiziert wird. Zugleich liegt kein Hinweis vor, dass abgesehen von den auftraggebenden Institutionen auch andere Beteiligte und Betroffene des untersuchten Programms in die Planung der Evaluation hinsichtlich von Informationsbedürfnissen einbezogen worden wären..	3	
N4 Auswahl und Umfang der Informationen	Konkrete Evaluationsfragestellungen zur Verfolgung der allgemein-abstrakt angegebenen Evaluationszwecke werden nicht angegeben. Ein Kontrollgruppen-Design wird angewendet. Dies erscheint sehr gut geeignet, die Frage des erreichten Impacts zu beleuchten, wenn es auch international unter EvaluationsexpertInnen anhaltende Fachdebatten über die Vor- und Nachteile des Kontrollgruppen-Designs gibt. Dem grundsätzlichen Nachteil des Kontrollgruppen-Ansatzes, dass es die Art und Weise der Erreichung der beobachtbaren Effekte im Dunkeln lässt („black box-Evaluation“), wird hier ein Stück weit gegengesteuert, indem die Datenerhebung auch auf Des Weiteren wird eine ergänzende Methode zur Anwendung gebracht, die zur Beleuchtung von Gesichtspunkten, die nicht nur erreichte Wirkungen betreffen, geeignet ist. Es wird in der Berichterstattung jedoch nicht erkennbar, dass Daten aus dieser Methode auch bei der Analyse und Einschätzung des Programms genutzt worden wären. Auf Ergebnisse einer früheren Evaluation desselben Programms und auf wissenschaftliche Literatur zum Themenbereich wird stellenweise Bezug genommen. Aspekte der Programmgestaltung, die in einer früheren Evaluation des untersuchten Programms als kritisch ausgewiesen wurden, werden gezielt überprüft. Das Programm wird nur in abrisshafter Form und hinsichtlich seiner beiden Hauptzielsetzungen greifbar gemacht. Eine Logic Chart oder eine ähnliche gesamthaft-strukturierende Aufarbeitung von Aufbau und Wirkungslogik des Programms wurde nicht erstellt. Die konzise methodische Vorgehensweise der Evaluation stößt in	2	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	der Umsetzung allerdings an Grenzen: Im Umgang mit den objektiven Daten, die aus ihrer Natur heraus keine Auskunft darüber geben, durch welche (intervenierenden) Faktoren die beobachteten Ergebnisse mit bedingt worden sein könnten, werden stellenweise spekulative Überlegungen angestellt, warum bestimmte Verhaltensweisen der Kontrollgruppe eingetreten sind. Zumindest einige der interessierenden Gesichtspunkte hätten in der Ausgestaltung des Erhebungsverfahrens wohl einbezogen werden können, um auch hierzu eine tragfähigere Datenlage zu erzeugen. Während die herangezogenen Daten aufgrund des gewählten Ansatzes kaum Aussagen über Kausalbeziehungen erlauben, was die EvaluatorsInnen auch klar festhalten, werden doch Annahmen über Kausalverhältnisse berichtet, die die EvaluatorsInnen aus eigenen Überlegungen treffen. Eine Heranziehung weiterer Informationsquellen hätte die Evaluation in dieser Hinsicht zusätzlich stützen können. Ergänzende Daten hätten unter Umständen auch auf weitere Faktoren im Programmgesehen hinweisen können, die in den herangezogenen Daten nicht erkennbar werden und in den von den EvaluatorsInnen angestellten Überlegungen keine Rolle spielen, sodass die Analyse der Entstehung von Programmwirkungen weiter vertieft werden hätte können. Der Berichtsabschnitt für AdressatsInnen auf europäischer Ebene ist in Entsprechung zu einem vorgegebenen Raster ausgearbeitet, zu dem kaum Gestaltungsmöglichkeiten vorgelegen sein dürften.	
N5 Transparenz von Werten	Die Evaluation ist am Vergleich mit einer Kontrollgruppe hinsichtlich der Erreichung von Programmzielen ausgerichtet. Im ausdrücklich als Überprüfung der Zielerreichung ausgewiesenen Berichtsteil werden allerdings nur wenige Daten unter Bezugnahme auf die beiden Hauptzielsetzungen des Programms herangezogen. Für weite Berichtsteile bleibt unklar, inwiefern sie nicht als solche ausgewiesene Bestandteile des Zielsystems des untersuchten Programms betreffen, oder eine ergänzende Analyse darstellen. Da keine umfassende Programmdarstellung angeboten wird, bleibt unklar, welcher Stellenwert diesen Beobachtungen für die Einschätzung des Programms zuzumessen ist. Zur Erreichung der zentralen Schlussfolgerungen werden jedoch auch solche Beobachtungen herangezogen. Es werden die Datenlagen in gleichsam selbstverständlicher Weise herangezogen, als ob sie automatisch aus sich selbst sprechen würden. Es werden keine Kriterien dargelegt bzw. formuliert, die den konkreten Zugang dazu festlegen würden, wie die beobachtete Datenlage zu interpretieren sein wird, um dem Programm mehr oder weniger Wert zuzumessen. Im Zusammenhang damit, dass Evaluationsfragestellungen nicht angegeben sind, erscheinen an manchen Stellen die Interpretationen der Datenlage nicht zwingend, sondern auch prinzipiell anders möglich. Auf derartige Möglichkeiten, unter Umständen punktuell auch eine andere Sichtweise auf das Programm zu gewinnen, oder Gründe, warum die Interpretation in der vorliegenden Form vorzuziehen ist, wird nicht eingegangen. Sichtweisen der Zielgruppen des Programms kommen insofern zum tragen, als ihre Auskünfte zu Gesichtspunkten der Zielerreichung einfließen. Es handelt sich jedoch nicht um einen Ansatz, der sich auf Werte stützt, die die vom Programm Betroffenen im Themenbereich vertreten, bzw. der solche Werte reflektiert.	3
N6 Vollständigkeit und Klarheit der Berichterstattung	Der Bericht ist klar strukturiert und enthält zahlreiche Tabellen und Grafiken, die auch dank einer sehr klaren Betitelung die Vermittlung wesentlicher Informationen gut unterstützen. Tabellen- und Abbildungsverzeichnisse sind beigegeben, sodass das Auffinden von Information durch die BerichtsadressatsInnen gut ermöglicht wird. Die Berichtssprache ist klar und strikt faktenorientiert, ohne die BerichtsadressatsInnen mit Technizismen übermäßig zu belasten. Griffige Darstellungen im Bericht werden durch getrennte Anhänge ergänzt, die Methoden Aspekte und Daten näher darstellen. Die Ergebnisse für die beiden AdressatsInnengruppen werden getrennt aufbereitet.	2

	Der Umstand, dass das untersuchte Programm nur abrisshaft vorgestellt wird, verleiht dem Text allerdings eher den Charakter eines Insider-Berichts für Adressatenkreise, die mit dem Programm bereits vertraut sind. Eine explizite Darstellung von Evaluationsfragestellungen und ihrer Korrespondenz zum Zielsystem bzw. der angenommenen Wirklogik des Programms hätte wesentlich dazu beitragen können, dass AdressatInnen dem Bericht mit seinen Argumentationsweisen und Schlussfolgerungen besser folgen können. Der Bericht enthält klar erkennbare und leicht auffindbare Kapitel zu den Schlussfolgerungen und Empfehlungen.	
N8 Nutzung und Nutzen der Evaluation	Aus den Angaben über einbezogene Beteiligte und Betroffene (vgl. N1) kann geschlossen werden, dass die Evaluation neben ihren Auftraggebern auch einige Stakeholder im Operationsbereich des Programms angesprochen hat, sowie zahlreiche begünstigte und nicht-begünstigte Mitglieder der Zielgruppe. Interesse an den Evaluationsergebnissen kann insbesondere im Bereich der in einen Workshop einbezogenen Stakeholder entstanden sein, wobei der genaue Einzugsbereich auf Grund fehlender Angaben jedoch im Unklaren bleibt. Weiterführende Angaben, ob und wie im Evaluationsprozess versucht wurde, Adressatenkreise für die Evaluation und ihre Ergebnisse zu interessieren und die spätere Nutzung vorzubereiten, finden sich nicht. Aus dem Grad der Einhaltung der in der Analyse herangezogenen Nützlichkeits-, Fairness- und Genauigkeitsstandards resultiert, dass dem Evaluationsbericht hohe Nützlichkeit zukommt, wenn auch eine perfekte Einlösung aller Standards nicht gegeben ist.	2
F3 Vollständige und faire Überprüfung	Der Bericht macht sowohl Stärken als auch Schwächen des untersuchten Programms explizit. Auch im Executive Summary werden identifizierte Stärken und Schwächen klar benannt, sodass auch eilige LeserInnen jedenfalls über sie informiert werden. Für weite Bereiche der Untersuchung erklären die EvaluatorenInnen nicht, ob die dargestellten Beobachtungen aus ihrer Sicht mit der Art und Weise zu tun haben, wie das Programm seine Ziele erreicht, ob es sich um nicht näher ausgewiesene Zielsetzungen handelt, oder ob Nebeneffekte beobachtet wurden. An mehreren Berichtsstellen wird klar darauf hingewiesen, dass die Daten aufgrund des gewählten Ansatzes kaum Aussagen über Kausalbeziehungen erlauben. Es werden gleichwohl Annahmen über Kausalverhältnisse berichtet, wobei die Herkunft dieser Annahmen nicht ausgewiesen wird. Die Tatsache, dass es sich um Überlegungen der EvaluatorenInnen und nicht um unmittelbare Ergebnisse handelt, ist durch die sprachliche Darstellungsweise erkennbar. Stellenweise erscheint die Interpretation von Differenzen zwischen ProgrammteilnehmerInnen und Kontrollgruppe defensiv, indem aus Grafiken entnehmbare Differenzen textlich nicht aufgegriffen werden. Dies kann unter Umständen dazu führen, dass verschiedene LeserInnen unterschiedliche Eindrücke von Aspekten bekommen, die mit Stärken und Schwächen zu tun haben können, je nachdem wie sie den Bericht lesen und verstehen. Eine Durchsicht der im Anhang beigegebenen Fragebögen zeigt, dass nicht über alle erhobenen Daten berichtet wird. Der Bericht macht keine Angaben, dass und warum dem so ist.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können	3
G2 Kontextanalyse	Die Evaluation unterhält eine explizite Beziehung zu der Tatsache, dass das Programm Effekte in einem Bereich erzielen will, den es selbst nicht direkt beeinflussen kann. Sie beleuchtet in ihrer Datenerhebung neben unmittelbaren Programmfakten auch erzielte Effekte im Kontext. Des Weiteren werden kontextuelle Bedingungen dafür, dass das Programm erwartete Effekte zeitigt, identifiziert.	3

	Informationsmaterialien zum Kontext werden herangezogen. Nicht an allen Analysepunkten, wo es von der Sache her möglich gewesen wäre, wird allerdings die mögliche Wirksamkeit von Kontextfaktoren reflektiert. Wenig intensiv eingegangen wird auf den Umstand, dass das untersuchte Programm auch durch thematisch verwandte Förderprogramme kontextualisiert ist. Es kann nicht von einer vollkommenen Vernachlässigung dieses Gesichtspunkts gesprochen werden, doch hätte eine intensivere Einlassung im Rahmen des Kontrollgruppen-Designs wohl auch stattfinden und so unter Umständen nähere Ergebnisse über die Positionierung und den Beitrag des Programms erbringen können. FördernehmerInnen der Kontextprogramme wurden nicht gezielt untersucht bzw. dem Kontrollgruppen-Design zugrunde gelegt.	
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht enthält ein eigenes Methodenkapitel, das die eingesetzten Methoden in zentralen Eckpunkten beschreibt und gemeinsam mit ausführlichen Methodendarstellungen im Anhang erlaubt, die Tragfähigkeit der mit ihrer Hilfe erhaltenen Informationen einzuschätzen. Methodische Angaben werden auch im direkten Konnex mit der Präsentation von Einzelergebnissen gemacht, was eine gute Auffassung der dargestellten Datenlagen ermöglicht. Ein klarer Bezug zwischen den verschiedenen erhobenen Fakten (Fragestellungen in der Umfrage bzw. Indikatoren) und der vollständigen Programmanlage mit allen Zielsetzungen wird allerdings nicht hergestellt. Dies kann für einen Bericht, der sich an die AuftraggeberInnen richtet, die über die Zwecke und Vorgehensweisen der Evaluation bereits informiert sind, als adäquat gelten. Für andere AdressatInnen bleibt der Zusammenhang jedoch schlecht erkennbar. Der Einsatz einer Methode wird zwar erwähnt, doch nicht hinsichtlich eines Stellenwerts in der Analysestrategie expliziert, Dort eventuell gewonnene Daten fließen nicht in erkennbarer Weise in die Berichterstattung ein, ohne dass dies kommentiert würde. In einem Anhang beantwortet die Evaluation gezielt Fragestellungen, die offenbar aus einem parallelen und eigens formulierten Evaluationsauftrag stammen. Es wird jedoch nicht erklärt, welche Vorgehensweisen in Bezug auf welchen Evaluationsauftrag gewählt wurden, die beide in einem Bericht abgehandelt werden, und auch nicht, weshalb kein Bezug zwischen Ergebnissen aus den beiden Teilen hergestellt wird. Ein wichtiger methodischer Aspekt des angewendeten Kontrollgruppen-Designs besteht in der Sicherheit, mit der Differenzen zwischen Treatment Group und Kontrollgruppe nachgewiesen werden können. Der Bericht bezieht sich an vielen Stellen auf die statistische Signifikanz der Differenzen, jedoch nicht überall, ohne dies und die Gründe dafür explizit zu machen.	3
G8 Begründete Schlussfolgerungen	Schlussfolgerungen und Empfehlungen werden in klarer Trennung von Daten und Dateninterpretationen dargestellt. Die hier vorgebrachten Schlussfolgerungen sind durch berichtete Daten gut gestützt, und die Empfehlungen werden einer klaren und strukturierten Argumentationsweise abgegeben.. Schlussfolgerungen werden auch im Berichtsverlauf im direkten Anschluss an Datenpräsentationen gezogen, was hohe Transparenz der Schlussfolgerungen im Verhältnis zur Datenlage gewährleistet. In einige dieser Schlussfolgerungen fließen Überlegungen der EvaluatorInnen ein, die sich nicht auf die sonst den Bericht tragenden exakten Daten stützen können. Dabei werden keine Überlegungen dargestellt, inwiefern auch andere Schlussfolgerungen aus der Datenlage möglich wären, sofern andere Annahmen getroffen werden. Da zu einigen eingesetzten Indikatoren nicht vollkommen transparent gemacht wird, inwiefern sie dem originären Zielsystem des Programms entstammen, oder ob eventuell für Zwecke der Evaluation auch Anpassungen vorgenommen wurden, können die aus der Datenlage gezogenen Schlussfolgerungen über den Wert des Programms in den Augen mancher BerichtsadressatInnen in ihrer Überzeugungskraft geschwächt sein.	2

Evaluationsbericht Nr. 20		Beobachtungszeitraum: 2011 - 2014	
Klassifikation	Einordnung		
Evaluationsrolle und -typ	summativ (Endbericht einer Begleitevaluation)		
Evaluationsschwerpunkte	Effektivität, Impact		
Standard	Kommentar	Einstufung*	
N1 Identifizierung der Beteiligten und Betroffenen	Im Rahmen der Datenerhebung wurden neben Programmverantwortlichen auch Stakeholder (Betroffene) in großer Zahl eingebunden. Aus der dem Bericht beigegebenen Liste der InterviewpartnerInnen ergibt sich eine breite Streuung auf alle relevanten Institutionen, wobei zumindest zum Teil die mittleren und gehobenen Steuerungsebenen angesprochen wurden. Angaben über eine Einbeziehung von Stakeholdern im Governance-System, für die die Evaluationsergebnisse Bedeutung haben könnten, werden nicht gemacht, und auch keine Angaben über die Interaktion mit den AuftraggeberInnen während des Evaluationsprozesses und etwaige Resultate daraus für die Durchführung der Evaluation. Die beiden vorangegangenen Zwischenberichte geben dazu ebenfalls keine Auskunft.	2	
N2 Klärung der Evaluationszwecke	Als Hauptzweck angegeben wird: Rahmenbedingungen für Akzeptanz und Wirkungsweise des Steuerungsmechanismus aufzeigen und die Erfahrungen für künftige Programme nutzbar machen. Untersuchung der Erreichung von Programmzielen. Es liegt eine klare Lernorientierung vor. Die Untersuchung von Zielerreichungen ist allerdings in die Frage der Wirkungsweise, der sie wohl prinzipiell zugerechnet werden könnte, nicht unmittelbar integriert. Es liegt kein Hinweis vor, dass abgesehen von den direkten AuftraggeberInnen andere Beteiligte und Betroffene des untersuchten Programms in eine Klärung der Evaluationszwecke einbezogen worden wären.	2	
N4 Auswahl und Umfang der Informationen	Zur Überprüfung von Zielerreichung wird eine adäquate Methode eingesetzt. Für das Erkennen der genaueren Wirkungsweise des Programms, wo ausdrücklich auch indirekte Wirkungen erfasst werden sollen, werden gut geeignete Methoden eingesetzt. In der Aufarbeitung der Programmkonzeption wird auf zentrale Annahmen eingegangen, wie das Programm seine intendierten Wirkungen erreichen wollte. Leitmotive der Analyse werden aus internationaler Fachliteratur bezogen. Während die Erbringung neuer aktueller Daten im Vordergrund steht, werden an verschiedenen Stellen auch Ergebnisse aus vorangegangenen Untersuchungen des Programms aufgegriffen. Dabei werden alle für die vorliegende Untersuchung relevanten Fakten nochmals dargestellt. Streckenweise werden in der Auseinandersetzung mit Zielerreichung Daten und Fragestellungen verfolgt, die im Verhältnis zu bereits erbrachten Ergebnissen kaum Mehrwert bringen. In der Auseinandersetzung mit der Wirkungsweise bleibt eine der eingesetzten Methoden mit ihren Teilergebnissen unverbunden stehen und wird in der Generierung der Gesamtergebnisse der Evaluation letztlich nicht genutzt. Die verfügbaren Strukturdaten decken die Programmlaufzeit nicht vollkommen ab, dies wird in der Ergebnisinterpretation jedoch nicht ausdrücklich berücksichtigt. Beim vergleichenden Umgang mit verschiedenen Strukturdaten wird nicht an allen Stellen völlig klar, inwieweit die aus unterschiedlichen Systematiken heraus entstandenen Daten den Untersuchungsgegenstand zur Gänze abdecken oder nur ausschnittsweise abbilden. Hier setzen sich offenbar durch die EvaluatorInnen festgestellte Unklarheiten auch in der Programmformulierung in der Analyse fort. Es kann anhand des vorliegenden Berichts nicht entschieden werden, ob eine gebotene Auseinandersetzung mit der Datenbasis nicht ausreichend geführt wurde, oder ob es sich um eine Schwäche der	2	

* 1. Standard sehr gut erfüllt, 2. Standard gut erfüllt, 3. Standard teilweise erfüllt, 4. Standard ansatzweise berücksichtigt, 5. Standard nicht erfüllt. 0. unzureichende Information

	Darstellungsweise handelt (vgl. F3). Ebenso wenig wird erkennbar, ob hierbei Informationsbedürfnisse des Auftraggebers abgeklärt wurden oder sogar gezielt befriedigt wurden (N1). Hinsichtlich der Generierung von „lessons learned“ hätte eine weitere Systematisierung von Verhaltensweisen der Zielgruppe und strukturelle Veränderungen in ihren Kontexten, unter Umständen unter Heranziehung weiterer ergänzender Informationen, produktiv sein können. Hier hätten auch Indikatoren verfolgt werden können, die positive indirekte Wirkungen des Programms abbilden.	
N5 Transparenz von Werten	Die Evaluation gibt an, dass einerseits an einer objektiven Überprüfung der Zielerreichung des untersuchten Programms orientiert ist, und andererseits an der Identifikation von direkten und indirekten Wirkungsweisen des Programms, wofür sie sich auf Sichtweisen der Betroffenen (Zielgruppe) stützt. Durch das eingesetzte Verfahren zur Datengenerierung und -analyse über die Wirkungsweise des Programms werden die Sichtweisen zentraler Betroffener in transparenter Weise zu einem tragenden Element der Analyse. Hinsichtlich der Zielerreichung wird dargestellt, dass sich in den Programmdokumenten zwei unterschiedliche Zieldefinitionen finden. Beide auf dieser Basis möglichen Sichtweisen auf den Evaluationsgegenstand werden verfolgt und gleichgewichtig nebeneinander dargestellt. Die konkreten Vorgehensweisen (Indikatorenwahl, subjektive Daten aus der Zielgruppe) werden allerdings ohne Erläuterung eingesetzt, als ob sie sich von selbst verstünden.	1
N6 Vollständigkeit und Klarheit der Berichterstattung	Die Berichterstattung erfolgt in einer Sprache, wie sie für Diplomarbeiten oder Dissertationen typisch ist. Während dies einerseits differenzierte Darstellungsweisen und Erläuterungen mit sich bringt, so leidet andererseits die rasche Auffassbarkeit des Berichts. An manchen Stellen wird aus der Formulierungsweise heraus nicht leicht erkennbar, welches Argument die EvaluatorenInnen hier vorbringen, bzw. wie es die Dateninterpretation beeinflusst. Ein Executive Summary ist nicht beigegeben. Wesentliche quantitative Ergebnisse werden teils in Grafiken und Tabellen dargestellt, die anhand der beigegebenen Verzeichnisse leicht aufgefunden werden können. Es werden allerdings an verschiedenen Stellen Daten rein textlich referiert, oder Verweise auf Tabellen in anderen Berichtsstellen angebracht. Nach dem ersten Berichtsdrittel werden keine Grafiken mehr eingesetzt, was die Auffassbarkeit der teils langwierig berichteten Ergebnisse deutlich erhöhen und dem Text Konzentration auf zentrale Aussagen ermöglichen hätte können. Die Präsentation und Interpretation qualitativer Daten ist flüssig und griffig, zahlreiche Zitate sorgen für gute Nachvollziehbarkeit. Das untersuchte Programm wird bereits in der Einleitung mit seiner Zielsetzung und den eingesetzten Mitteln dargestellt. Die Programm-konzeption wird ausführlich beschrieben, wobei auch auf während der Programmlaufzeit erfolgte Konzeptveränderungen eingegangen wird. Essentielle Merkmale des Kontexts, in dem die Zielgruppe relevante Verhaltensweisen an den Tag legt, werden ebenfalls dargestellt. Die Anschlussfähigkeit für die einbezogenen Stakeholder (vgl. N1) und weitere Interessierte kann durch mehrere Darstellungsdefizite im Bereich der quantitativen Methode als beeinträchtigt gelten.	2
N8 Nutzung und Nutzen der Evaluation	Hinweise zu einbezogenen Beteiligten und Betroffenen des Programms (vgl. N1) deuten darauf hin, dass eine große Zahl von Personen im Operationsbereich des Programms auf die Evaluation aufmerksam wurde, sodass sie sich in der Folge für deren Ergebnisse interessieren konnte. Weitere Angaben über etwaige Schritte, die Interesse und Aufgreifen von Ergebnissen bei den HauptadressatInnen und potenziellen weiteren Adressatenkreisen befördern hätten können, finden sich nicht. Es kann davon ausgegangen werden, dass der Bericht seine AdressatInnen dabei unterstützt, neue Sichtweisen auf das Programm zu entwickeln und Steuerungsinstrumente in der Zukunft sinnvoll anzupassen. Schwächen in den Standards N6, F3,	2

	G3 und G8 schränken die Nützlichkeit allerdings ein.	
F3 Vollständige und faire Überprüfung	Die Berichterstattung erfolgt in einer neutral gehaltenen, faktenbezogenen Sprache. Es wird festgehalten, dass die Evaluation ihre Aufgabe vor allem darin sieht, hinderliche Faktoren für die Zielerreichung durch das Steuerungsinstrument zu identifizieren, um Lernen für zukünftige Programme zu ermöglichen. Der Bericht nimmt die beabsichtigte Überprüfung der Zielerreichung in strikt neutraler Form vor. Analysen zur Wirkungsweise des Programms und „lessons learned“ konzentrieren sich auf die Identifikation von Schwächen, wodurch in Entsprechung zur Zielsetzung der Evaluation diesbezügliche Schlussfolgerungen und Empfehlungen erbracht werden können. Die Analyse der Zielerreichung kann als vollständig gelten insofern, als alle verfügbaren amtlichen Daten herangezogen wurden. Jedoch wird nicht klar diskutiert, welche Stärken und Schwächen bestimmte Datenquellen für die Beantwortbarkeit einer der Hauptfragestellungen der Evaluation haben. Es kann davon ausgegangen werden, dass die AuftraggeberInnen des Berichts über die Fachkenntnisse verfügen, die eine irrtumsfreie Auffassung der aus Strukturdaten gewonnenen Ergebnisse ermöglichen. Für nicht mit der Evaluation und ihrem Gegenstand Vertraute hätte es hier aber klarerer Hinweise und Explikationen bedurft. Die Relation von Programmbeteiligung und von insgesamt eingetretenen Entwicklungen in der Zielgruppe, die im Sinne der Programmzielsetzung sind, wird nicht interpretiert. Der Beitrag des Programms zu den insgesamt eingetretenen Entwicklungen wird somit von den EvaluatorenInnen nicht eingeschätzt.	3
F5 Offenlegung der Ergebnisse	Es liegt ein publizierter Bericht vor, der von einer Website abgerufen werden kann, die den InteressentInnen allerdings vorab bekannt sein muss. Abgesehen davon liegt nur unzureichende Information vor, um die Erfüllung des Standards einschätzen zu können.	3
G2 Kontextanalyse	Das untersuchte Programm zielt auf Veränderungen in einem Umfeld ab, das es nur bedingt und indirekt beeinflussen kann. Die Evaluation setzt gezielt an dieser Konfiguration an und formuliert die passende Fragestellung, inwiefern das Steuerungsinstrument unter dort anzutreffenden Handlungsbedingungen als zielführend eingeschätzt werden kann. Zur Beantwortung werden Daten zu Merkmalen und Entwicklungen des Kontextes herangezogen. Einsatzpunkte anderer Steuerungsinstrumente, durch die während der Programmlaufzeit synergetisch auf die Erreichung der Programmziele hingewirkt werden sollte, werden berücksichtigt und auch in den Empfehlungen angesprochen.	1
G3 Beschreibung von Zwecken und Vorgehen	Der Bericht enthält kein eigenes Methodenkapitel. Genauere Evaluationsfragen werden nicht dargestellt. Die angegebenen Evaluationszwecke und eingesetzten Methoden stehen jedoch in einem transparenten, gut nachvollziehbaren Verhältnis, und in direktem Bezug zur Programmdarstellung. Der Bericht gibt genaue und klare methodische Hinweise zu allen verwendeten Daten. Diese finden sich im Berichtstext im direkten Konnex zu präsentierten Ergebnissen. In einem Anhang werden die InterviewpartnerInnen gelistet. Die herangezogenen Strukturdaten decken die Programmlaufzeit fast vollständig, aber doch nicht zur Gänze ab, dies wird in der Ergebnisinterpretation jedoch nicht explizit benannt. Es finden sich Passagen, die eine vollkommene Abdeckung der Programmlaufzeit suggerieren, und gleichzeitig - seltener - solche, die von einer vorläufigen Einschätzung sprechen.	2
G8 Begründete Schlussfolgerungen	Ausgesprochene Schlussfolgerungen und Empfehlungen haben klaren Rückhalt in präsentierten Faktenlagen. Schlussfolgerungen werden auch im Zuge der Dateninterpretation im Text gezogen, was ihre Nachvollziehbarkeit erleichtert, aber doch auch ihre Auffindbarkeit behindert. Die Schlussfolgerung zum Ausmaß der Zielerreichung ist wohlbegründet und transparent. Verbesserungsvorschläge, die von Betroffenen des Programms erhalten wurde, werden als solche berichtet und nicht mit Schlussfolgerungen der EvaluatorenInnen vermischt.	2

	In einem abschließenden Kapitel werden zentrale Erkenntnisse zur Wirkungsweise des Programms zusammengefasst. Allerdings werden dabei nicht alle Daten zusammengezogen, aus erhobenen und berichteten Daten wird nicht die volle Konsequenz gezogen. Punktuell kommt es zu leichten Umgewichtungen gegenüber den zuvor berichteten Dateninterpretationen und Schlussfolgerungen, da nicht alles, was an relevanten Gesichtspunkten zuvor berichtet wurde, aufgegriffen wird. Überraschender Weise sind Aussagen über den Grad der Zielerreichung hier nicht enthalten. Stattdessen werden nochmals Berechnungen präsentiert, die für die Beantwortung der Hauptfragestellungen der Evaluation nicht relevant sind. Im letzten Kapitel werden Empfehlungen abgegeben, die teils spezifisch und umsetzbar sind, sich teils aber auch in der Feststellung von virulenten Problemen erschöpfen.	
--	---	--

Anhang 3: Sampling der Evaluationsberichte

Grundgesamtheit der Programmevaluationen mit Publikation im Zeitraum 1.1.2003 - 30.9.2014				
Programmevaluationen nach Zeitabschnitten				
	2011-2014	2007-2010	2003-2006	Total
absolut	16	9	21	46
%	34,8	19,5	45,7	100
Programmevaluationen nach Auftraggebern				
% der Evaluationen	2011-2014	2007-2010	2003-2006	Total
BMVIT und Vorgängerinstitutionen	15,1	9,4	13,2	37,7
BMWF und Vorgängerinstitutionen	9,4	3,8	5,7	18,9
BMWFJ und Vorgängerinstitutionen	11,3	3,8	11,3	26,4
Agenturen (AWS, FFG, FWF)	3,8	1,9	11,3	17,0
gesamt	39,6	18,9	41,5	100
% der Auftraggeberschaften	2011-2014	2007-2010	2003-2006	Total
BMVIT und Vorgängerinstitutionen	17,4	10,9	15,2	43,5
BMWF und Vorgängerinstitutionen	10,9	4,3	6,5	21,7
BMWFJ und Vorgängerinstitutionen	13,0	4,3	13,0	30,4
Agenturen (AWS, FFG, FWF)	4,3	2,2	13,0	19,6
gesamt	45,7	21,7	47,8	115,2
Programmevaluationen nach Evaluationsinstituten				
% der Evaluationen	2011-2014	2007-2010	2003-2006	Total
KMU Forschung Austria	4,8	3,2	6,3	14,3
Technopolis Group	7,9	4,8	9,5	22,2
andere Österreich	12,7	4,8	17,5	34,9
andere Ausland	7,9	4,8	15,9	28,6
gesamt	33,3	17,5	49,2	100,0
% aller Tätigkeiten von Evaluationsinstituten	2011-2014	2007-2010	2003-2006	Total
KMU Forschung Austria	6,5	4,3	8,7	19,6
Technopolis Group	10,9	6,5	13,0	30,4
andere Österreich	17,4	6,5	23,9	47,8
andere Ausland	10,9	6,5	21,7	39,1
gesamt	45,7	23,9	67,4	137,0

Samplingkriterien zur Stichprobenziehung				
Sample von 20 Programmevaluationen nach gewichteten Zeitabschnitten				
	2011-2014	2007-2010	2003-2006	Total
Sample in % nach Gewichtung	50,0	25,0	25,0	100
Gewichte für Zeitabschnitte	1,437	1,282	0,547	-
Sample in % der Grundgesamtheit	0,435	0,435	0,435	-
Sample absolut	10	5	5	20
Programmevaluationen nach Auftraggebern				
Auftraggeberschaften, absolut (Richtwert-Obergrenzen)				
	2011-2014	2007-2010	2003-2006	Total
BMVIT und Vorgängerinstitutionen	5,0	2,8	1,7	9,5
BMWF und Vorgängerinstitutionen	3,1	1,1	0,7	5,0
BMWFJ und Vorgängerinstitutionen	3,8	1,1	1,4	6,3
Agenturen (aws, FFG, FWF)	1,3	0,6	1,4	3,2
gesamt	13,1	5,6	5,2	23,9
Evaluationen, absolut (Richtwert-Untergrenzen)				
	2011-2014	2007-2010	2003-2006	Total
BMVIT und Vorgängerinstitutionen	4,3	2,0	2,0	8,3
BMWF und Vorgängerinstitutionen	2,0	1,0	0,6	3,6
BMWFJ und Vorgängerinstitutionen	3,3	1,0	1,2	5,5
Agenturen (AWS, FFG, FWF)	1,1	1,0	1,2	3,3
gesamt	10,0	5,0	5,0	20,0
Programmevaluationen nach Evaluationsinstituten				
Tätigkeiten von Evaluationsinstituten, absolut (Richtwert-Obergrenzen)				
	2011-2014	2007-2010	2003-2006	Total
KMU Forschung Austria	1,9	1,1	1,0	3,9
Technopolis Group	3,1	1,7	1,4	6,2
andere Österreich	5,0	1,7	2,6	9,3
andere Ausland	3,1	1,7	2,4	7,2
gesamt	13,1	6,1	7,3	26,6
Evaluationen, absolut (Richtwert-Untergrenzen)				
	2011-2014	2007-2010	2003-2006	Total
KMU Forschung Austria	1,4	0,9	0,6	3,0
Technopolis Group	2,4	1,4	1,0	4,7
andere Österreich	3,8	1,4	1,8	7,0
andere Ausland	2,4	1,4	1,6	5,4
gesamt	10,0	5,0	5,0	20,0

Erreichte Sampling-Quoten in der gezogenen Stichprobe				
Programmevaluationen nach Auftraggebern				
	2011-2014	2007-2010	2003-2006	Total
BMVIT und Vorgängerinstitutionen	5	2	2	9
BMWF und Vorgängerinstitutionen	3	1	1	5
BMWFJ und Vorgängerinstitutionen	4	1	1	6
Agenturen (aws, FFG, FWF)	1	1	1	3
gesamt	13	5	5	23
Programmevaluationen nach Evaluationsinstituten				
	2011-2014	2007-2010	2003-2006	Total
KMU Forschung Austria	1	1	1	3
Technopolis Group	3	1	2	6
andere Österreich	5	2	2	9
andere Ausland	4	1	2	7
gesamt	13	5	7	25

ANHANG 4: Herangezogene DeGEval-Standards mit Interpretations- hintergrund der Joint Committee-Standards

In der folgenden Aufstellung wird zunächst jeder DeGEval-Standard mit Nummer, Titel, Kurzbeschreibung und erläuterndem Text angegeben. Im Anschluss daran wird jeweils – grau hinterlegt - der korrespondierende, zur genaueren Interpretation angeratene JC-Standard (Handbuch der Evaluationsstandards, Joint Committee & Sanders 2006) mit Nummer, Titel und Kurzbeschreibung zitiert.

Eine vollständige Darstellung aller Angaben zu jedem JC-Standard würde den Rahmen dieses Berichts jedenfalls sprengen. Es werden im Folgenden auszugsweise erläuternde Texte, Richtlinien und Häufige Fehler angeführt, sofern sie im Lauf der durchgeführten Berichtsanalyse zu Klärungen beigetragen haben bzw. im Umgang mit den Berichten schlagend wurden.

Die Wiedergabe beider Standards erfolgt auf Basis der in den beiden Standard-Publikationen erteilten Freigabe zur Verwendung im fachlichen Kontext.

N1 Identifizierung der Beteiligten und Betroffenen
Die am Evaluationsgegenstand beteiligten oder von ihm betroffenen Personen bzw. Personen- gruppen sollen identifiziert werden, damit deren Interessen geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können. Diese Akteure, die im Umfeld einer Evaluation und ihres Gegenstandes berücksichtigt werden sollen, können durch Fragen wie die folgenden erschlossen werden: • Wer hat über die Zukunft des Evaluationsgegenstandes zu entscheiden? • Wer ist verantwortlich für die Konzeption und die Gestaltung des Evaluationsgegenstandes? • Wer ist an der praktischen Umsetzung des zu untersuchenden Gegenstandes beteiligt? • Wer soll durch den Evaluationsgegenstand direkt oder indirekt erreicht werden (Zielgruppen und deren soziales Umfeld)? Die so identifizierten Personen, Gruppen und Organisationen werden als „Beteiligte und Betroffene“ bezeichnet. Es soll dabei auch festgestellt werden, wem durch den zu evaluierenden Gegenstand Nachteile erwachsen können oder wer z. B. von einem Programm ausgeschlossen wird. Soweit dies die finanziellen und zeitlichen Bedingungen zulassen, sollen die identifizierten Beteiligten und Betroffenen den Evaluationszwecken entsprechend in die Planung und Durchführung der Evaluation einbezogen werden (Heiner 1998). Es ist besonders wichtig, die Informationsbedürfnisse der Adressatinnen und Adressaten der Evaluation und ihrer Ergebnisse zu ermitteln, sie bei der Klärung der Zwecksetzungen und Fragestellungen zu beteiligen und die Evaluation entsprechend auszurichten. Dabei sollen auch weitere an den Evaluationsergebnissen interessierte Gruppen berücksichtigt werden (z. B. Entscheidungsträger, welche ähnliche Projekte planen, Fachöffentlichkeit, allgemeine Öffentlichkeit). Eine Evaluation, welche die Beteiligten und Betroffenen mit gestalten und die sich nach deren Informationsbedürfnissen ausrichtet, hat beste Aussichten, von den Adressaten und Adressatinnen beachtet und tatsächlich genutzt zu werden.
JC - N1 Ermittlung der Beteiligten und Betroffenen Die an einer Evaluation beteiligten oder von ihr betroffenen Personen sollten identifiziert werden, damit deren Interessen und Bedürfnisse berücksichtigt werden können. (...) Zentral ist, dass potentiell wichtige Beteiligte & Betroffene identifiziert und soweit dies die finanziellen und zeitlichen Einschränkungen zulassen, in die Planung und Durchführung der Evaluation miteinbezogen werden. (...) Findet die Identifikation der Beteiligten & Betroffenen nicht statt, könnte die Evaluation zu einer fehlgeleiteten akademischen Übung werden, deren Ergebnisse kritisiert, ignoriert oder abgelehnt würden, da sie keinerlei spezifische Fragestellung der Beteiligten & Betroffenen behandeln. (...) Richtlinie A: Personen in führenden Positionen sollten zuerst kontaktiert werden, da sie der Evaluation helfen können, weitere Beteiligte & Betroffene zu ermitteln. Richtlinie B: Es sollten Vertreter der identifizierten Beteiligten & Betroffenen kontaktiert werden, um

herauszufinden, welche Bedeutung diese der Evaluation beimessen, wie sie die Ergebnisse verwenden möchten und welche bestimmten Informationen nützlich wären. Wo nötig, sollten sie auch dabei unterstützt werden, realistische Erwartungen an die Evaluation zu richten, die den gegebenen methodischen, finanziellen und politischen Einschränkungen entsprechen.

Richtlinie D: Mit dem Auftraggeber sollte eine Übereinkunft hinsichtlich der relativen Bedeutung der potentiellen Beteiligten & Betroffenen und deren Informationswünschen erzielt werden, und die Datenerhebung sowie die Berichterstattung sollte entsprechend geplant und umgesetzt werden.

Richtlinie E: Während der gesamten Evaluation sollte man aufmerksam auf weitere Beteiligte & Betroffene achten, denen die Evaluation von Nutzen sein sollte; und man sollte sich die Flexibilität und Fähigkeit bewahren, im Rahmen der gegebenen zeitlichen und finanziellen Grenzen auf ihre Bedürfnisse einzugehen.

Häufiger Fehler A: Es wird zugelassen, dass der Auftraggeber den Kontakt der Evaluatorin zu anderen Beteiligten & Betroffenen in unangemessener Weise einschränkt.

Häufiger Fehler B: Es wird stillschweigend unterstellt, allen Informationswünschen der Beteiligten und Betroffenen zu entsprechen, wenn dies in Wirklichkeit nicht möglich ist.

Häufiger Fehler C: Es wird davon ausgegangen, dass Personen in Führungs- oder Entscheidungspositionen die einzigen oder die wichtigsten Beteiligten und Betroffenen sind.

Häufiger Fehler E: Es wird nicht zwischen Auftraggebern und anderen Beteiligten und Betroffenen unterschieden.

N2 Klärung der Evaluationszwecke

Es soll deutlich bestimmt sein, welche Zwecke mit der Evaluation verfolgt werden, so dass die Beteiligten und Betroffenen Position dazu beziehen können und das Evaluationsteam einen klaren Arbeitsauftrag verfolgen kann.

Die Zwecke einer Evaluation sollen frühzeitig angesprochen und im Verlauf einer Evaluation ausgehandelt, schriftlich fixiert und so weit wie möglich den Beteiligten und Betroffenen bekannt gemacht werden, damit diese sich in die Evaluationsplanung einbringen können. Die Tätigkeit der beauftragenden Person bzw. Institution und des Evaluationsteams wird durch klare Zwecksetzungen erleichtert, da sie damit eine gute Grundlage haben, um die Verbreitung der Evaluationsergebnisse und deren Nutzung vorzubereiten.

Für die Klärung der Zwecksetzungen ist die idealtypische Unterscheidung dreier Hauptzwecke von Evaluationen hilfreich (Patton 1997, S. 79):

- Informationen bereit zu stellen, die geeignet sind, der schrittweisen Gestaltung des Evaluationsgegenstandes, z. B. bezüglich der Optimierung von Konzepten oder Prozessen, Orientierung zu geben.
- Informationen bereit zu stellen, die geeignet sind, einer grundlegenden Entscheidung zum Evaluationsgegenstand Orientierung zu geben, z. B. Einführung, Weiterführung, Ausweitung oder Einstellung eines Programms betreffend.
- Erkenntnisse bereit zu stellen, welche der öffentlichen und der politischen oder der wissenschaftlichen Diskussion Anregungen geben können.

Steht bei den ersten beiden Hauptzwecken der instrumentelle Nutzen im Vordergrund, so ist mit dem Erkenntnisgewinn der konzeptionelle Nutzen angesprochen (Cronbach u.a. 1980, S. 112 ff.). Die gleichzeitige und gleichrangige Verfolgung mehrerer Hauptzwecke kann die Nutzung der Evaluation beeinträchtigen. Es bietet sich daher oft an, eine klare Priorität zu setzen und verschiedene Hauptzwecke in zeitlich getrennten Phasen oder arbeitsteilig durch unterschiedliche Evaluationsteams zu bearbeiten.

keine Entsprechung

N4 Auswahl und Umfang der Informationen
Auswahl und Umfang der erfassten Informationen sollen die Behandlung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und gleichzeitig den Informationsbedarf des Auftraggebers und anderer Adressaten und Adressatinnen berücksichtigen.
Bei der Planung eines Evaluationsprojektes soll heraus gearbeitet werden, welche Informationen zur Beantwortung der Evaluationsfragestellungen unabdingbar sind und welche Informationen zwar interessant und wünschenswert, aber für die Beantwortung der Kernfragestellungen unerheblich sind. Es ist dabei darauf zu achten, dass die vorhandenen Ressourcen zur Datenerhebung entsprechend den Prioritäten zur Beantwortung der Kernfragestellungen und dem Informationsbedarf der wichtigsten vorgesehenen Nutzer und Nutzerinnen der Evaluation eingesetzt werden (Bundesamt für Gesundheit 1997).
JC - N3 Umfang und Auswahl der Informationen
Die gewonnenen Informationen sollten von einem Umfang und einer Auswahl sein, welche die Behandlung sachdienlicher Fragen zum Programm ermöglichen und gleichzeitig auf die Interessen und Bedürfnisse des Auftraggebers und anderer beteiligter & Betroffener eingehen.
Um hinreichend informativ zu sein, müssen Evaluationen für die Ziele der Entscheidungsträger relevant, für bedeutende Beteiligte & Betroffene wichtig und so umfassend sein, dass sie eine Beurteilung von Verwendbarkeit und Güte ermöglichen. Die meisten Evaluationen sind mit einer Vielzahl von Beteiligten & Betroffenen konfrontiert, die alle die Gelegenheit bekommen sollten, die Evaluation mitzugestalten, und die ein Recht auf Zugang zu den Ergebnissen haben sollten. Darüber hinaus sollten die Evaluatoreninnen jedoch bestrebt sein, alle wichtigen Dimensionen des Programms (z.B. Effektivität, unerwünschte Nebenwirkungen, Kosten, Bezugnahme auf die Bedürfnisse der Lernenden, Sinnhaftigkeit der dem Programm zugrundeliegenden Annahmen und Wertvorstellungen und Durchführbarkeit) zu bewerten, unabhängig davon, ob die Beteiligten & Betroffenen speziell um diese Informationen bitten. Es ist wichtig, das Programm zu untersuchen und zu dokumentieren, um sicherzustellen, dass eine geeignete Abgrenzung der Evaluation erfolgt. In der Regel ist es nicht möglich, alle Informationen zusammenzutragen, die zu Beginn als bedeutend erachtet werden. Für eine gute Evaluation sind jedoch nicht alle Informationen gleich wichtig und nicht alle sind unbedingt erforderlich. Sobald das Informationspotential bestimmt ist, sollte eine Selektion stattfinden, so daß auf Unwichtiges verzichtet und der Schwerpunkt auf das Wesentliche gelegt wird. Dieses Selektionsverfahren erfordert Urteilskraft. Die Evaluatoreninnen müssen bestimmen, was dem Auftraggeber wichtig erscheint, sie sollten jedoch auch auf bedeutende Bereiche hinweisen, die von diesem möglicherweise übersehen wurden (...). Weiterhin sollten die Evaluatoreninnen themenbezogene Literatur - darunter solche zu früheren Evaluationen, theoretische Abhandlungen und Forschungsberichte - und Diskussionsbeiträge der Beteiligten & Betroffenen konsultieren, um die zu behandelnden Fragestellungen herausfiltern und gewichten zu können. Es muss auch sorgfältig darauf geachtet werden, dass die Durchführung der Evaluation nicht von bereits festgelegten Evaluationsverfahren bestimmt wird - z.B. von der Forderung dass standardisierte Tests angewendet werden. (...) Vor Beginn der Datenerhebungen sollte der Evaluationsplan mit den vorgesehenen Informationserhebungen vom Auftraggeber und weiteren Beteiligten & Betroffenen eingesehen werden. Auf diese Weise kann sichergestellt werden, daß die zu erhebenden Informationen die wichtigen Themen behandeln. Die Evaluatorenin sollte auch beachten, dass während der Evaluation zusätzliche Fragestellungen und unerwartete Informationen auftauchen könnten, die dann eventuell in den Evaluationsprozess integriert werden müssen. Richtlinie A: Die Forderungen des Auftraggebers an die Evaluation müssen den Evaluatoreninnen klar sein. Richtlinie B: Vertreter wichtiger Beteiligter & Betroffener sollten interviewt werden, um einen Einblick in ihre unterschiedlichen und möglicherweise widersprüchlichen Standpunkte sowie ihre Informationsbedürfnisse zu gewinnen. Richtlinie C: Es sollte nicht der Eindruck vermittelt werden, dass alle Fragestellungen beantwortet werden können. Richtlinie D: Den Beteiligten & Betroffenen sollte geholfen werden, in Anbetracht der verfügbaren finanziellen, zeitlichen und personellen Mittel realistische Erwartungen zu entwickeln.

Richtlinie E: Der Auftraggeber sollte aufgefordert werden, die potenziellen Nutzer der Evaluation nach ihrer Bedeutung zu ordnen, und in Zusammenarbeit mit Vertretern aller B&B sollten auch die Themen nach ihrer Bedeutung für den jeweiligen Adressaten gewichtet werden.
Richtlinie F: Es sollte mit dem Auftraggeber zusammengearbeitet werden, um die von den verschiedenen Adressaten benannten und nach ihrer Bedeutung geordneten Themen zu sichten, unwichtigere Punkte von der Liste zu streichen und andere, welche die Evaluatoreninnen für wichtig erachten, hinzunehmen, auch wenn sie nicht nachgefragt wurden.
Richtlinie G: Es sollte genug Flexibilität vorhanden sein, um neue Fragestellungen und unvorhergesehene Informationen aufzugreifen, die sich im Lauf der Evaluation ergeben.
Richtlinie H: Während aller Evaluationsphasen (Datenbeschaffung, Auswertung, Interpretation und Berichterstattung) sollte sich die Arbeit an der endgültigen Themenliste orientieren, wobei den Punkten mit hoher Priorität die größte Beachtung zukommen sollte.
Häufiger Fehler A: Es wird nicht in jeder Phase der Evaluation (Evaluationsplanung, Budgetierung, Erhebung, Auswertung, Interpretation und Darstellung der Informationen) zwischen Ausführlichkeit und Selektivität abgewogen.
Häufiger Fehler C: Es wird versäumt, durch regelmäßige Kontakte mit dem Auftraggeber und anderen Beteiligten & Betroffenen die Informationserfordernisse zu aktualisieren.
Häufiger Fehler D: Die Informationsgewinnung richtet sich eher nach Bequemlichkeitskriterien (z.B. wenn bestimmte Instrumente bereits zur Verfügung stehen) als nach Notwendigkeitskriterien.
Häufiger Fehler E: Der Umfang der Informationen wird nicht eingegrenzt, d.h. es wird nicht festgelegt und in jeder Evaluationsphase berücksichtigt, welche Fragen behandelt werden und was das Ziel der Evaluation ist.
Häufiger Fehler F: Es werden Informationen gesammelt, die für das Hauptziel der Evaluation irrelevant sind.

N5 Transparenz von Werten

Die Perspektiven und Annahmen der Beteiligten und Betroffenen, auf denen die Evaluation und die Interpretation der Ergebnisse beruhen, sollen so beschrieben werden, dass die Grundlagen der Bewertungen klar ersichtlich sind.

In mehreren Phasen der Evaluation, bereits bei der Festlegung der Evaluationszwecke und -fragestellungen, spielen die Werte der Beteiligten und Betroffenen eine große Rolle. Deren Identifikation und Thematisierung durch die Evaluation ist dann besonders wichtig, wenn die Beteiligten und Betroffenen zur aktiven Teilnahme an der Evaluation und schließlich zur Nutzung der Evaluationsergebnisse ermuntert werden sollen (Wottawa/Thierau 1998). Die in der Schlussphase vorzunehmende Interpretation der gesammelten Informationen und Ergebnisse stellt einen der wichtigsten und kritischsten Punkte in einem Evaluationsprozess dar. In diesem Interpretationsvorgang spielen gesellschaftlich vermittelte Werthaltungen (Normen) unvermeidlich eine große Rolle. Um den Interpretationsvorgang überzeugend, nachvollziehbar und beurteilbar zu machen, ist es unerlässlich, die zu Grunde liegenden Werthaltungen möglichst transparent zu machen

JC - N4 Feststellung von Werten

Die Perspektiven, Verfahren und Gedankengänge, auf denen die Interpretationen der Ergebnisse beruhen, sollte sorgfältig beschrieben werden, damit die Grundlagen der Werturteile klar ersichtlich sind.

(...) Das Bewerten - die Einschätzung oder Klassifizierung einer Sache nach ihrer Nützlichkeit und ihrem allgemeinen Wert - ist die grundlegende Aufgabe jeder Evaluation. Im Mittelpunkt dieser Aufgabe steht das Erfordernis, die bei einer Evaluation gewonnenen Informationen zu interpretieren. Solche Informationen - ob quantitativ oder qualitativ, prozeß- oder produktbezogen, formativ oder summativ - sind nur von geringem Interesse oder Nutzen, wenn sie nicht anhand einer geeigneten und vertretbaren Idee dazu, was Wert hat und was nicht, interpretiert werden.

Die Interpretation der gewonnenen Informationen ist jedoch ein aufwendiges und umstrittenes Unterfangen. Es muss dabei unter anderem entschieden werden, wer die Werturteile vornimmt und welche Verfahren dazu verwendet werden. (...)

Der Kernpunkt dieses Standards ist, dass Evaluatoreninnen und ihre Auftraggeber zusammen mit den verschiedenen Beteiligten & Betroffenen sorgfältig festlegen sollten, welcher Ansatz zugrundegelegt werden soll, um den gewonnenen Informationen Wert zuzuweisen. Den gewählten Ansatz sollten sie dann offenlegen und begründen.

Richtlinie A: Es sollten verschiedene Grundlagen für die Interpretation der Ergebnisse in Betracht gezogen werden, z.B. Programmziele, Prozeßbeschreibungen, Gesetze und Verordnungen, institutionelle Leitziele, demokratische Ideale, gesellschaftliche Normen, Leistungen einer Vergleichsgruppe, der abgeschätzte Bedarf einer Kundengruppe, die erwartete Leistung der Stichproben-Gruppe, fachliche Standards und Urteile, die von verschiedenen Referenzgruppen geäußert werden.

Richtlinie B: Es sollte überlegt werden, wer Interpretationen vornehmen soll, z.B. die Evaluatoreninnen, der Auftraggeber, die verschiedenen Beteiligten & Betroffenen, eine Steuerungsgruppe oder eine Kombination von diesen.

Richtlinie C: Es sollten verschiedene Verfahren in Erwägung gezogen werden, mit denen man den gesammelten Informationen Wertbedeutungen zuschreiben kann. So könnten z.B. unterschiedliche Teams damit beauftragt werden, parteinehmende Berichte zu erstellen; das zu evaluierende Programm könnte von einer Jury beurteilt werden; oder es könnte versucht werden, mittels einer Delphi-Studie einen Konsens zu finden.

Richtlinie D: Wenn bei einer Evaluation verschiedene gleichermaßen vertretbare Wertvorstellungen vorliegen, sollte über diese Alternativen und ihre Vor- und Nachteile berichtet werden.

Häufiger Fehler A: Es wird angenommen, dass Evaluationen objektiv im Sinn von „wertfrei“ sein können.

Häufiger Fehler B: Es wird nicht bestimmt, welche Wertorientierungen (z.B. pädagogische, soziale, wirtschaftliche und wissenschaftliche) der Auftraggeber und die Beteiligten & Betroffenen bei der Interpretation der Ergebnisse für wichtig halten.

Häufiger Fehler C: Bei der Entwicklung von Verfahren zur Datenbeschaffung und -analyse wird nicht berücksichtigt, welche Kriterien - wie z.B. Leistungen einer Kontrollgruppe oder Leistungen nach Maßgabe eines vorher festgelegten Maßstabs - benötigt werden, um die Ergebnisse zu interpretieren.

N6 Vollständigkeit und Klarheit der Berichterstattung

Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen, leicht zu verstehen und nachvollziehbar sein.

Eine Bedingung für die erfolgreiche Kommunikation von Evaluationsergebnissen ist die Vollständigkeit und Klarheit des schriftlichen Berichts (oder anderer Formen der Berichterstattung). Die Sprache soll für die Adressaten und Adressatinnen verständlich sein und wichtige Begriffe sollen definiert und konsistent verwendet werden. Die Zusammenstellung der wichtigsten Ergebnisse – z. B. in einer voran gestellten Zusammenfassung, in Tabellenform oder in Graphiken – kann das Verständnis der berichteten Inhalte erheblich fördern.

Die Gestaltung und Art der Berichterstattung soll auf die Wahrnehmungsgewohnheiten der Adressaten und Adressatinnen einer Evaluation abgestimmt werden. Ein umfangreicher Schlussbericht stellt dabei nicht für jede adressierte Gruppe die optimale Form zur Weitergabe der Informationen dar. Zuweilen können – je nach Situation und Zielpublikum – Referate, Workshops oder ähnliche Kommunikationsformen zweckdienlicher sein.

JC - N5 Klarheit des Berichts

Evaluationsberichte sollen das evaluierte Programm einschließlich seines Kontextes ebenso beschreiben wie die Ziele, die Verfahren und Befunde der Evaluation, damit die wesentlichen Informationen zur Verfügung stehen und leicht verstanden werden können.

Es gibt viele verschiedene Arten von Berichten, darunter mündliche Rückmeldungen, schriftliche Mitteilungen und ausführliche Dokumente. Unabhängig vom verwendeten Medium (...) ist Klarheit unbedingt erforderlich, damit der Bericht verständlich ist, glaubwürdig wirkt und Verwendung findet. Um von Nutzen zu sein, muss eine Evaluation verstanden werden. Die Beteiligten & Betroffenen sollten leicht erkennen können, welche Ziele die Evaluation verfolgte, was evaluiert wurde, wie die Evaluation durchgeführt wurde, welche Informationen gewonnen wurden, zu welchen

Schlußfolgerungen man kam und – gegebenenfalls - welche Empfehlungen gemacht wurden. (...) Richtlinie A: Es sollte sorgfältig überlegt werden, welches Medium für die Berichterstattung an die vorgesehene Adressatengruppe geeignet ist. Richtlinie B: Die Präsentation des Berichts sollte so kurz, einfach und direkt wie möglich gehalten werden und sich auf die Behandlung der Fragestellungen der Evaluation konzentrieren. Richtlinie C: Es sollte erwogen werden, sowohl einen zusammenfassenden als auch einen technischen Bericht zu erstellen und dabei verschiedene Medien einzusetzen. Die Berichte sollten auf die unterschiedlichen Adressaten (Auftraggeber und Beteiligte & Betroffene) zugeschnitten sein. Richtlinie D: Es sollten ausreichende Informationen zum Kontext, zum Programm und zur Evaluation vermittelt werden, um eine gesicherte Grundlage für Schlussfolgerungen und Empfehlungen herzustellen (...). Richtlinie G: Zusammenfassende Aussagen sollten durch eine klare Diskussion zusammenhängender Probleme, Ziele, Fragestellungen und ggf. möglicher Bereiche zukünftiger Studien ergänzt werden (...). Häufiger Fehler A: Bei der Entscheidung über die Art der Berichterstattung wird nicht berücksichtigt, inwieweit bei den Beteiligten & Betroffenen Fachkenntnisse vorhanden sind. Häufiger Fehler C: Es wird nicht über unterschiedliche Sichtweisen berichtet, auch wenn dadurch Mehrdeutigkeit erzeugt würde. Häufiger Fehler E: Die Berichte werden nicht so gestaltet, dass sie leicht gelesen werden können. Häufiger Fehler H: Das Programm wird nicht genügend beschrieben, weil angenommen wird, daß alle Beteiligten & Betroffenen bereits über dessen Besonderheiten informiert sind.
--

N8 Nutzung und Nutzen der Evaluation
Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation aufmerksam zur Kenntnis zu nehmen und ihre Ergebnisse zu nutzen.
Selbst mit hoher Professionalität erzeugte Ergebnisse werden häufig nicht genutzt oder ganz anders, als dies in den Vereinbarungen zur Evaluation geplant war. Wichtige Voraussetzungen für eine intendierte Verwendung liegen in einer hohen Durchführbarkeit, Fairness und Genauigkeit von Evaluationen sowie den Aspekten, die in den vorausgegangenen sieben Nützlichkeitsstandards ausgeführt sind. Besonders die Umsetzung von Schlussfolgerungen oder Empfehlungen hängt stark davon ab, ob die Adressaten und Adressatinnen die Erwartung haben, dass das Evaluationsprojekt ihnen und ihren Vorhaben von Nutzen ist. Dabei ist es nicht selten so, dass unterschiedliche Adressaten auch unterschiedliche Nutzenerwartungen haben, die sich widersprechen oder sich gegenseitig ausschließen können. Eine wichtige Voraussetzung, um eine vereinbarte bzw. intendierte Nutzung und damit auch den Nutzen von Evaluationen zu fördern, ist die angemessene Einbeziehung der verschiedenen Adressaten und Adressatinnen in die Planung und die Durchführung der Evaluation sowie die Aufbereitung von Ergebnissen. Ebenfalls positiv wirkt es sich aus, wenn im Verlauf der Untersuchung kontinuierlich und transparent Rückmeldungen gegeben werden (Stockbauer 2000). Dies fördert, dass abgesehen vom Ergebnisnutzen bereits im Verlauf einer Evaluation ein Prozessnutzen eintritt (Patton 1998).
JC - N7 Wirkung der Evaluation Evaluationen sollen so geplant, durchgeführt und dargestellt werden, dass die Beteiligten & Betroffenen dazu ermuntert werden, dem Evaluationsprozess zu folgen, damit die Wahrscheinlichkeit steigt, dass die Evaluation genutzt wird.
Der Begriff Wirkung (<i>impact</i>) einer Evaluation bezieht sich darauf, welchen Einfluß die Evaluation auf die Entscheidungen und das weitere Vorgehen der Adressaten hat. Er bezieht sich außerdem auf den konzeptionellen Einfluß der Evaluation auf die Beteiligten & Betroffenen. Die EvaluatorenInnen sollten die Beteiligten & Betroffenen dabei unterstützen, die Evaluationsergebnisse zu nutzen, um sinnvolle Veränderungen vorzunehmen, wie z.B. Verbesserung der Programme, (...). Des Weiteren sollten die EvaluatorenInnen den Beteiligten & Betroffenen helfen, durch die Evaluation neue Perspektiven zum Programm zu entwickeln. Die EvaluatorenInnen dürfen jedoch nicht davon ausgehen, dass Verbesserungen automatisch

stattfinden, sobald der Evaluationsbericht fertiggestellt ist. Tatsächlich ist es so, dass die Evaluatorinnen möglicherweise eine der folgenden Rollen spielen müssen: die Evaluationstätigkeiten so zu planen, mit Personal zu besetzen und durchzuführen, dass den Adressaten geholfen wird, die Ergebnisse zu bewerten und konstruktiv zu nutzen; oder: den Beteiligten & Betroffenen neue Sichtweisen zu eröffnen. Die Evaluatorinnen sollten jedoch nicht in die Rolle des Auftraggebers schlüpfen.

Richtlinie A: Den hauptsächlich Beteiligten & Betroffenen sollte zu Beginn der Evaluation erklärt werden, in welcher Weise die Ergebnisse ihrer Arbeit von Nutzen sein können.

Richtlinie C: Die Berichterstattung an die B&B sollte offen, aufrichtig und konkret sein, und für eventuelle Rückfragen sollten die Evaluatorinnen bereitwillig zur Verfügung stehen (siehe N5, Klarheit des Berichts, und K6, Offenlegung der Ergebnisse).

Richtlinie D: In regelmäßigen Abständen sollte über Zwischenergebnisse berichtet und speziell darauf hingewiesen werden, welche Bedeutung diese für die verschiedenen Beteiligten & Betroffenen haben.

Richtlinie E: Die Vorzüge von sinnvollen Handlungsalternativen sollten mit den B&B bewertet und im Schlußbericht behandelt werden.

Richtlinie F: Die schriftlichen Berichte sollten durch mündliche Mitteilungen laufend ergänzt werden.

Richtlinie G: Soweit Zeit, finanzielle Mittel und personelle Ressourcen es erlauben, sollten die Evaluatorinnen den Beteiligten & Betroffenen helfen, nach Freigabe des Schlußberichts die Evaluationsergebnisse zu bewerten, zu interpretieren und umzusetzen.

Häufiger Fehler E: Es wird angenommen, dass die Auftraggeber alles lesen, was man ihnen vorlegt.

Häufiger Fehler F: Es wird nicht eingegriffen, wenn bemerkt wird, dass Evaluationsergebnisse mißbraucht oder falsch interpretiert werden.

Häufiger Fehler G: Die verschiedenen Nutzer der Evaluation werden nicht individuell angesprochen.

F3 Vollständige und faire Überprüfung

Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst vollständig und fair überprüfen und darstellen, so dass die Stärken weiter ausgebaut und die Schwachpunkte behandelt werden können.

Es ist selbstverständlich, dass die Schwächen eines Evaluationsgegenstandes umfassend überprüft und dargestellt werden sollen. Nur so können sie behoben werden. Doch auch die Darstellung der Stärken ist wichtig. Beides ist für die umfassende Bewertung des Evaluationsgegenstandes und für seine Verbesserung notwendig.

Stärken und Schwächen eines Evaluationsgegenstandes hängen oft eng zusammen. Daher ist zu beachten, dass unter Umständen durch die Korrektur von Schwächen bestehende Stärken gefährdet werden können. Andererseits ist es eventuell möglich, bestehende Schwächen durch vorhandene Stärken auszugleichen. Häufig ist nicht nur der beabsichtigte Abbau von Schwächen, sondern auch der Ausbau von Stärken Anlass einer Evaluation. Darüber hinaus steigt in der Regel die Bereitschaft zur Mitarbeit und die Akzeptanz der Evaluation durch die Beteiligten und Betroffenen, wenn auch die Stärken erfasst werden.

Die Überprüfung und Darstellung der Stärken und Schwächen soll so vollständig wie möglich sein, damit alle wichtigen Aspekte zum Vorschein kommen. Sie soll gegenüber den Beteiligten und Betroffenen fair sein. Auftraggeber oder andere Beteiligte sollen nicht versuchen, einseitig Einfluss auf die Evaluation und den Bericht zu nehmen. Es ist Aufgabe der Evaluatoren und Evaluatorinnen, eine vollständige, ausgewogene und faire Evaluation sicherzustellen.

JC -K5 Vollständige und faire Einschätzung

Evaluationen sollen in der Überprüfung und in der Präsentation der Stärken und Schwächen des evaluierten Programms vollständig und fair sein, so dass die Stärken weiter ausgebaut und Problemfelder angesprochen werden können.

Für Ausgewogenheit bei einer Evaluation zu sorgen bedeutet nicht, dass Stärken und Schwächen in gleicher Anzahl herausgearbeitet werden sollen. Vielmehr ist damit gemeint, dass man vollständig und auf faire Art und Weise sowohl die negativen als auch die positiven Aspekte des zu bewertenden

Programms bewerten und im Bericht nennen soll.

Selbst wenn das vorrangige Ziel einer Evaluation darin besteht, die Schwächen eines Untersuchungsobjekts zu bestimmen, ist es notwendig, auch die Stärken festzustellen: Einerseits könne Stärken manchmal genutzt werden, um Schwächen zu korrigieren. Andererseits können durch Maßnahmen zur Behebung von Schwachpunkten auf unbeabsichtigte Weise unerkannte Stärken verringert werden.

Die Methoden, die in einer Evaluation benutzt werden, sollten zusammen mit ihren Stärken und Schwächen von der Evaluatorin beschrieben werden, damit ihre Auswirkungen auf die Ergebnisse und Schlußfolgerungen bewertet werden können.

Richtlinie A: Es sollte vollständig über Ergebnisse berichtet werden, die auf Stärken wie auf Schwächen hinweisen, seien sie nun beabsichtigt oder unbeabsichtigt (...).

Richtlinie C: Wenn aufgrund zeitlicher oder finanzieller Beschränkungen einige für die Evaluation relevante Daten nicht zugänglich sind, sollte über diese Auslassungen berichtet werden, und die sich daraus ergebenden Auswirkungen auf die Gesamtberichterstattung des Programms sollten eingeschätzt werden.

Häufiger Fehler A: Die Berichterstattung über Stärken und Schwächen wird manipuliert, um Einzelpersonen oder Interessengruppen zufriedenzustellen oder es wird zugelassen, dass im Bericht Schwächen ausgeklammert werden, die sich als peinlich erweisen könnten, oder um die persönlichen Interessen und Neigungen der Evaluatorin zu fördern oder zu schützen.

Häufiger Fehler C: Bei der Beurteilung oder Interpretation der Ergebnisse werden diese entweder als Stärken oder als Schwächen dargestellt, ohne dass andere Sichtweisen in Betracht gezogen werden, die möglicherweise zu einer anderen Schlußfolgerung führen würden.

Häufiger Fehler D: Im Evaluationsbericht werden spekulative oder vorläufige Ergebnisse genannt, um für eine Ausgewogenheit von Stärken und Schwächen zu sorgen.

Häufiger Fehler E: Es wird nicht darüber berichtet, welche Methoden bei der Evaluation verwendet wurden und wie sich Stärken und Schwächen auf die Ergebnisse und Schlussfolgerungen ausgewirkt haben könnten.

F5 Offenlegung der Ergebnisse

Die Evaluationsergebnisse sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden.

Grundsätzlich sollen alle Personen, die an einer Evaluation beteiligt oder von dieser betroffen sind, Zugang zu den Ergebnissen und zu dem Evaluationsbericht haben. Alle an einer Evaluation mitwirkenden Akteure (z. B. Evaluationsteam, Auftraggeber, Adressaten und Adressatinnen, Mitarbeiter und Mitarbeiterinnen der Organisation, in der evaluiert wird) sind dafür verantwortlich, dass dies gewährleistet wird. Da der Kreis der Beteiligten und Betroffenen oft sehr groß ist, kann dies in vielen Fällen bedeuten, dass der Bericht öffentlich zugänglich gemacht werden muss.

Die Offenlegung der Evaluationsergebnisse soll den Nutzen der Evaluation steigern. Es kann jedoch Fälle geben, in denen eine vollständige Offenlegung den Nutzen einer Evaluation mindert. Dies kann z. B. der Fall sein, wenn konkurrierende Firmen die Evaluationsergebnisse zum Schaden einer evaluierten Firma nutzen können, wenn die Bereitschaft der Beteiligten und Betroffenen zur Mitarbeit an der Evaluation und der Umsetzung ihrer Ergebnisse durch ihr Wissen um die spätere Veröffentlichung (von Schwächen) leidet, oder wenn die Gefahr besteht, dass Informationsgeber durch eine Veröffentlichung kompromittiert würden. In solchen Fällen ist es die Aufgabe der Akteure, gemeinsam eine Lösung zu finden und ggf. eine Beschränkung der Offenlegung zu vereinbaren. Beschränkungen bei der Offenlegung sollen begründet werden.

Die Offenlegung der Evaluationsergebnisse, ihre Art und ihr Umfang, soll zu Beginn einer Evaluation vereinbart und vertraglich festgelegt werden. Falls dies nicht möglich ist, soll zu Beginn vereinbart werden, durch wen, wie, nach welchen Kriterien und wann (z. B. erst zum Ende eines Evaluationsprojektes) über die Offenlegung der Ergebnisse entschieden wird.

JC - K6 Offenlegung der Ergebnisse

Die Vertragsparteien einer Evaluation sollten sicherstellen, dass die Evaluationsergebnisse - einschließlich ihrer relevanten Beschränkungen - den durch die Evaluation betroffenen Personen ebenso wie all jenen, die einen ausgewiesenen Anspruch auf die Evaluationsergebnisse haben,

zugänglich gemacht werden.

Personen, die von der Evaluation betroffen sind, sowie Personen und Einrichtungen, die per Gesetz ein Recht auf die Ergebnisse der Evaluation haben, sollten vollständig über das Wie und Warum einer Evaluation und über die Gesamtheit der Ergebnisse informiert werden, es sei denn, dass die Offenlegung solcher Informationen die öffentliche Sicherheit gefährdet oder persönliche Freiheiten beschneidet.

Die Beteiligten & Betroffenen einer Evaluation haben einen Anspruch auf Berichte, in denen die Urteile und Empfehlungen der Evaluatoreninnen sowie die Informationen und Methoden, die zu deren Formulierung verwendet wurden, klar und deutlich ausgeführt sind. Offenlegung von Ergebnissen bedeutet, dass so aufrichtig wie möglich mitgeteilt wird, was man als Evaluatorsin denkt und glaubt, wobei die Beurteilung bei bestmöglicher Informiertheit vorgenommen werden sollte.

Wenn Personen oder Gruppen, die von der Evaluation betroffen sein werden, keine Informationen darüber erhalten können, ist es auch nicht möglich, dass sie Mängel bei Verfahren und Daten aufdecken. Ebenso wenig können sie die Ergebnisse konstruktiv nutzen. (...) Von einer Evaluation sollte erwartet werden können, dass sie der kritischen Überprüfung durch Personen standhält, deren Leben sie möglicherweise beeinflusst, und sie sollte diese mit für sie nützlichen Informationen versorgen.

Die Fähigkeit der Evaluatorsin, alle von der Evaluation betroffenen Personen zu identifizieren und ihnen gerecht zu werden, hat bedeutenden Einfluss auf die Fairness und Nützlichkeit der Evaluation. Die Möglichkeit der Evaluatorsin, Informationen freizugeben oder zu veröffentlichen, unterliegt häufig der teilweisen Kontrolle des Auftraggebers. Daher liegt die Verantwortung zur Einhaltung dieses Standards sowohl bei der Evaluatorsin als auch beim Auftraggeber. Dieser Punkt sollte als Teil der formalen Vereinbarung ausgehandelt werden.

Richtlinie B: Es sollte ein vollständiger schriftlicher und wenn möglich auch mündlicher Bericht über die Evaluation abgegeben werden, der relevante Ergebnisse ohne Auslassungen oder sonstige Abänderungen offenlegt.

Richtlinie C: Die Grundlage für die Zusammenhänge zwischen dem Zweck der Evaluation, den verwendeten Methoden, den erhobenen Daten und den Ergebnissen sollte klar herausgestellt werden (siehe G 10, Begründete Schlussfolgerungen).

Richtlinie D: Es sollen relevante Sichtweisen sowohl von Befürwortern als auch von Kritikern des zu evaluierenden Programms präsentiert werden.

Richtlinie E: Der Bericht sollte Urteile und Empfehlungen enthalten, die für eine breite, ausgewogene und auf Informationen basierende Sichtweise stehen.

Richtlinie F: Es sollte über Schlüsselfaktoren berichtet werden, welche die Vertretbarkeit der Evaluation deutlich vermindern oder steigern könnten - egal ob sie vor oder während der Evaluation entdeckt wurden - und ihre Auswirkungen auf die Ergebnisse und Empfehlungen der Evaluation sollten offen diskutiert werden.

Richtlinie G: Soweit geeignet, sollten zu Entwicklungs- und Umsetzungsphasen Zwischenberichte erstellt werden, die dazu beitragen, notwendige Änderungen im Programm herbeizuführen.

Richtlinie H: Die Auftraggeber sollten aufgefordert werden, alle betroffenen Personen angemessen, rechtzeitig und in geeigneter sprachlicher Form mit Informationen zu versorgen, was dazu beiträgt, dass diese sich zu aufgeklärten Befürwortern, Verbrauchern, Kritikern und Beobachtern entwickeln.

Häufiger Fehler A: Bei der Feststellung der Adressatengruppen für die Evaluationsberichte geht man so vor, wie es am bequemsten und sparsamsten ist, statt sich von ethischen und rechtlichen Überlegungen leiten zu lassen.

Häufiger Fehler B: Die Evaluatorsinnen sind nicht an der Überwachung und Freigabe von Informationen beteiligt, welche die Evaluation betreffen oder sich aus dieser ergeben.

Häufiger Fehler C: Dem Auftraggeber wird erlaubt, Teile des Evaluationsberichts ohne Rücksprache mit der Evaluatorsin auszuwählen und freizugeben.

Häufiger Fehler D: Dem Auftraggeber wird die einseitige Befugnis erteilt, den Evaluationsbericht vor seiner Freigabe zu bearbeiten, zu zensieren oder auf andere Weise zu ändern.

Häufiger Fehler E: Ausgewählte Informationen werden nur einigen Mitgliedern einer Gruppe mitgeteilt, die von der Evaluation betroffen ist, während sie den übrigen Mitgliedern vorenthalten werden.

Häufiger Fehler F: Es werden Berichte herausgegeben, die gezielt abgeändert wurden, um den Eigeninteressen der Evaluatorsin, des Auftraggebers oder der Programmmitarbeiter zu dienen.

G2 Kontextanalyse
Der Kontext des Evaluationsgegenstandes soll ausreichend detailliert untersucht und analysiert werden.
Den Kontext des Evaluationsgegenstandes bilden die Rahmenbedingungen, welche diesen Evaluationsgegenstand umgeben. Dies sind z. B. das soziale und politische Klima, die Charakteristika und Interessen der maßgeblich Beteiligten und Betroffenen oder die wirtschaftlichen Rahmenbedingungen. Analysen zum Umfeld des Evaluationsgegenstandes sind bedeutsam zur Gewinnung von Erkenntnissen über mögliche Bedingungsfaktoren, die sich auf den Evaluationsgegenstand und auf die Evaluationsergebnisse auswirken. Eine fundierte Kontextanalyse erlaubt es auch, die Resultate einer Evaluation in Bezug auf ihre Übertragbarkeit auf andere Kontexte einschätzen zu können. Es soll vermieden werden, den Kontext zu eng zu definieren. Es ist aber auch davor zu warnen, die Kontextanalyse zu ausführlich anzulegen, da dies auf Kosten der Bearbeitung der anderen Evaluationsschritte gehen kann.
JC - G2 Kontextanalyse
Der Kontext, in dem das Programm angesiedelt ist, soll ausreichend detailliert untersucht werden, damit mögliche Beeinflussungen des Programms identifiziert werden können.
Zum Kontext des Programms zählen die geographische Verortung des Programms, der zeitliche Rahmen, das politische und soziale Umfeld, parallel laufende konkurrierende Aktivitäten, das Personal und relevante wirtschaftliche Bedingungen. Diese und andere Kontextfaktoren müssen bei der Planung, bei der Durchführung und beim Bericht der Evaluation untersucht, beschrieben und berücksichtigt werden. Jedes Programm ist in seinen Kontext eingebettet und auf schwer durchschaubare Weise damit verwoben. Kenntnisse des Umfelds sind erforderlich, wenn die Evaluation realistisch sein und auf die Bedingungen, in denen sich das Programm abspielt, eingehen soll. (...)

G3 Beschreibung von Zwecken und Vorgehen
Gegenstand, Zwecke, Fragestellungen und Vorgehen der Evaluation, einschließlich der angewandten Methoden, sollen genau dokumentiert und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.
Um den Evaluationsprozess transparent zu gestalten, sind die Zwecke, die mit einer Evaluation angestrebt werden, schon im Verlauf des Evaluationsprozesses genau zu dokumentieren und im Rahmen der Berichterstattung an die Adressatinnen und Adressaten der Evaluation in nachvollziehbarer Weise zu kommunizieren. Ebenso soll mit den Fragestellungen, die bearbeitet werden, sowie den gewählten Vorgehensweisen, Methoden und den Begründungen für die dem Evaluationsplan zu Grunde liegenden Entscheidungen verfahren werden. Der Zeitplan und Abweichungen sind schriftlich festzuhalten.
Bei der Beschreibung der Zwecke und der Fragestellungen der Evaluation ist speziell darauf zu achten, dass auch abweichende Sichtweisen berücksichtigt werden. Die Dokumentation des Vorgehens umfasst die detaillierte Beschreibung der organisatorischen Vorkehrungen, der Erhebungen einschließlich evtl. Stichprobenverfahren, der Aufbereitung und Auswertung der Daten, des Interpretationsprozesses und schließlich der Berichterstattung. Bei der Vorgehensweise ist darauf zu achten, dass sich diese im Verlauf des Evaluationsprozesses verändern kann und dass sich damit das geplante vom tatsächlichen Vorgehen unterscheiden kann. Es soll auch nach außen deutlich dargelegt werden, welche Abweichungen entstanden sind und welche Gründe dafür verantwortlich sind.
G3 Beschreibung von Zwecken und Vorgehen
Die Zwecke und das Vorgehen der Evaluation sollten ausreichend genau dokumentiert und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.
Die Zwecke einer Evaluation werden durch die einzelnen Evaluationsziele und die vorgesehenen Nutzungen der Evaluationsergebnisse definiert. (...) Die Verfahren einer Evaluation umfassen die

Mittel und Wege, wie Daten und Informationen gesammelt, organisiert, analysiert und dargestellt werden, um dem Zweck einer Evaluation gerecht zu werden oder ihn zu erfüllen. (...) Auch im Verlauf einer Evaluation sollten die Zwecke und Verfahren in regelmäßigen Abständen überprüft werden, um ein realistisches Bild davon zu bekommen, ob die ursprünglichen Evaluationspläne geeignet waren, inwieweit sie umgesetzt wurden und in welchem Maße Änderungen stattfanden. Beim Abschluss der Evaluation benötigen die Nutzer eine aktuelle Beschreibung davon, was die Evaluatoreninnen und mit welcher Absicht tatsächlich gemacht haben, damit die Ergebnisse und Empfehlungen entsprechend beurteilt und interpretiert werden können.

Solche Zweck- und Verfahrensbeschreibungen dienen auch weiteren Vorhaben: Sie stellen unerlässliche Informationen für jeden dar, der die Evaluation evaluieren will. Sie können Informationen für die Durchführung von ähnlichen Evaluationen unter anderen Rahmenbedingungen liefern. Sie legen die Zwecke und Verfahren fest, auf die in Wiederholungen der Evaluation geachtet werden soll. Außerdem können die Beschreibungen von Evaluationszwecken und -verfahren gute Fallbeispiele sein, die in der Ausbildung von Evaluatoreninnen verwendet werden können.

Richtlinie E: Größere Veränderungen bei den Evaluationszwecken und -verfahren sollten mit dem Auftraggeber klar abgesprochen sein.

Richtlinie H: Die Zwecke und Verfahren sollten zum Abschluss der Evaluation sowohl in einem zusammenfassenden Bericht (Bericht an Entscheidungsträger) als auch in einem vollständigen technischen Bericht beschrieben werden. Abweichungen von ursprünglichen Plänen sollten bezeichnet werden (siehe N5, Klarheit des Berichts).

Häufiger Fehler A: Es wird angenommen, dass eine Evaluationsofferte eine vollständige Beschreibung der Zwecke und Verfahren einer abgeschlossenen Evaluation enthält.

Häufiger Fehler B: Es wird angenommen, dass die Zwecke und Verfahren, die der Auftraggeber und die Evaluatorin zu Beginn der Evaluation vereinbart haben, sich im Laufe der Evaluation nicht mehr verändern.

Häufiger Fehler C: Es wird nicht vorgesehen, Anpassungen der Zwecke und Verfahren vorzunehmen, die sich im Verlauf der Evaluation als notwendig erweisen könnten.

Häufiger Fehler D: Allein die Tatsache, dass die Zwecke und Verfahren sorgfältig beschrieben sind, führt zu der Schlussfolgerung, dass sie angemessen und gut sind.

G8 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen sollen ausdrücklich begründet werden, damit die Adressaten und Adressatinnen diese einschätzen können.

Die Schlussfolgerungen einer Evaluation sollen explizit begründet und zusammen mit den zugrunde liegenden Annahmen und den eingesetzten Evaluationsverfahren transparent dargestellt werden. Dabei ist auch deutlich auf den Geltungsbereich der Schlussfolgerungen hinzuweisen. Wo erforderlich, soll die Berichterstattung eine Diskussion alternativer Interpretationen beinhalten, wobei begründet werden soll, warum diese verworfen wurden. Die Einhaltung dieses Standards erlaubt es den Nutzerinnen und Nutzern der Evaluationsergebnisse, die Aussagekraft der Folgerungen einzuschätzen. Zudem wird die Überzeugungskraft der Folgerungen gestärkt.

JC - G10 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen sollten ausdrücklich begründet werden, damit die Beteiligten & Betroffenen diese einschätzen können.

Die aus einer Evaluation gezogenen Folgerungen, die Beurteilungen und Empfehlungen darstellen, müssen begründet sein und müssen auch tatsächlich gerechtfertigt werden. Wenn die Beteiligten & Betroffenen keine ausreichenden Informationen erhalten, um festzustellen, ob die Schlussfolgerungen berechtigt sind, mißachten sie diese womöglich. Außerdem sind Schlussfolgerungen mit unzureichender Absicherung möglicherweise nicht ganz korrekt, wodurch die Adressaten zu unangebrachten Tätigkeiten verleitet werden können. Schlussfolgerungen müssen sich auf alle gesammelten Informationen stützen, die in der Sache relevant sind. Sie müssen die Ergebnisse gründlicher Analyse und logischer Überlegungen enthalten. Es sollten vollständige Informationen darüber mitgeliefert werden, wie die Evaluation durchgeführt wurde (d.h. ein Bericht über die Verfahren, Informationen und zugrundeliegenden Annahmen der Evaluation). Zusammen mit den

Schlussfolgerungen sollten möglichst auch andere plausible Erklärungen für die Ergebnisse erörtert werden und es sollte dargelegt werden, warum diese Erklärungen verworfen wurden.

R-A: Es sollten Schlussfolgerungen entwickelt werden, die sowohl die Fragestellungen der Adressaten beantworten als auch die Verfahren und Ergebnisse der Evaluation wahrheitsgetreu widerspiegeln.

R-B: Informationen, die mit den Schlussfolgerungen verbunden sind, sollten dargelegt werden.

R-C: Es sollten plausible andere Erklärungen der Ergebnisse erarbeitet, bewertet und im Bericht dargelegt werden. Wenn möglich sollte erklärt werden, warum diese Erklärungen nicht berücksichtigt werden sollten.

R-D: Schlussfolgerungen sollten auf die Situationen, Zeiträume, Personen, Zusammenhänge und Zwecke beschränkt werden, für welche die Evaluationsergebnisse anwendbar sind.

R-E: Den Adressaten sollte empfohlen werden, bei der Interpretation von mehrdeutigen Ergebnissen aus dem Evaluationsbericht vorsichtig zu sein..

Häufiger Fehler A: Die Evaluatoreninnen konzentrieren sich darauf, die Fragestellungen der Adressaten zu beantworten, ohne die Begrenzungen der Evaluationsverfahren und -daten ausreichend zu berücksichtigen.

Häufiger Fehler B: Bei der Erarbeitung von Schlussfolgerungen über die Wirksamkeit des Programms werden mögliche Nebenwirkungen außer acht gelassen.

Häufiger Fehler C: Die Schlussfolgerungen stützen sich auf unzureichende oder ungenaue Informationen.

Häufiger Fehler D: Bei der Interpretation der Evaluationsergebnisse wird zu vorsichtig vorgegangen.

Häufiger Fehler E: Die Beschränkungen der Evaluationsstudie werden nicht dargelegt.

Anhang 5: Evaluationsschwerpunkte nach OECD DAC Standards¹

RELEVANCE

The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.

In evaluating the relevance of a programme or a project, it is useful to consider the following questions:

- To what extent are the objectives of the programme still valid?
- Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives?
- Are the activities and outputs of the programme consistent with the intended impacts and effects?

EFFECTIVENESS

A measure of the extent to which an aid activity attains its objectives.

In evaluating the effectiveness of a programme or a project, it is useful to consider the following questions:

- To what extent were the objectives achieved / are likely to be achieved?
- What were the major factors influencing the achievement or non-achievement of the objectives?

EFFICIENCY

Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs. It is an economic term which is used to assess the extent to which aid uses the least costly resources possible in order to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been adopted.

When evaluating development co-operation programmes and projects it is useful to consider the following criteria, laid out in the DAC Principles for Evaluation of Development Assistance.

When evaluating the efficiency of a programme or a project, it is useful to consider the following questions:

- Were activities cost-efficient?
- Were objectives achieved on time?
- Was the programme or project implemented in the most efficient way compared to alternatives?

IMPACT

The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results and must also include the positive and negative impact of external factors, such as changes in terms of trade and financial conditions.

When evaluating the impact of a programme or a project, it is useful to consider the following questions:

- What has happened as a result of the programme or project and why?
- What real difference has the activity made to the beneficiaries?

¹ Quelle: OECD DAC Network on Development Co-Operation (2010), Summary of Key Norms and Standards, 2nd ed., Part II.

SUSTAINABILITY

Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable.

When evaluating the sustainability of a programme or a project, it is useful to consider the following questions:

- To what extent did the benefits of a programme or project continue after donor funding ceased?
- What were the major factors which influenced the achievement or non-achievement of sustainability of the programme or project?

ANHANG 6: Liste der InterviewpartnerInnen

Name	Institution / Themenbereich	Interviewdatum
Mag. Michael Binder	FFG, Leiter Stabstelle Strategie und Datenanalyse	30.6.2015
Univ.-Prof. Dr. Günther R. Burkert	BMWWFV Sektion V, AL V/8 Forschungspolitik von Universitäten, Fachhochschulen und Privatuniversitäten	8.7.2015
Dr. Leonhard Jörg	FFG, Stabstelle Strategie und Datenanalyse	30.6.2015
Mag. Norbert Knoll, MSc.	AWS, Strategie und Evaluierung	30.7.2015
MinR Mag. Josef Mandl	BMWWFV, AL C1/11 Innovation und Transfer	21.7.2015
Mag. Simone Mesner	BMWWFV Sektion V, AL V/1 Forschung, Technologie, Innovation: Strategie, Berichtswesen, Wissenstransfer aus Universitäten und öffentlichen Forschungseinrichtungen.	20.7.2015
Mag.Dr. Rupert Pichler	BMVIT Sektion III, AL Abt. I 2 – Forschungs- und Technologieförderung	7.7.2015
Mag. Sabine Pohoryles-Drexel	BMWWFV Abt. C1/12 FTI- Strategie und Internationale Forschungs- und Technologie-kooperationen	9.7.2015
Dr. Falk Reckling	FWF, Abteilungsleitung Strategie-Analysen	30.7.2015
Mag. Ingolf Schädler	BMVIT Sektion III – Innovation und Telekommunikation, Leiter Bereich Innovation	7.7.2015
Ing. Mag. Karl Schiller	AWS, Leiter Hochtechnologieförderung	30.7.2015
Dr. Michael Stampfer	Geschichte und Struktur der Plattform fteval (WWTF, Geschäftsführer)	22.7.2015
SC Mag. Barbara Weitgruber, MA	BMWWFV SL Sektion V Wissenschaftliche Forschung; Internationale Angelegenheiten	24.7.2015
Mag. Michael Wiesmüller	BMVIT Sektion III, AL Abt. I 5 – Informations- und industrielle Technologien, Raumfahrt	31.7.2015
Dr. Klaus Zinöcker	FWF, Abteilung Strategie-Analysen	30.7.2015

ANHANG 7: Interviewleitfaden

1. Wenn Sie auf die Evaluationstätigkeit Ihrer Organisation über die letzten zehn Jahre zurück blicken: Wie wurden die Ergebnisse der verschiedenen Evaluationen (Daten, Schlussfolgerungen, Empfehlungen) genutzt? Und inwiefern haben die Evaluationen auch in anderer Weise Nutzen gebracht?
2. Was waren im Rückblick Eigenschaften einzelner Evaluationen, durch die sie für Ihre Organisation nützlicher oder weniger nützlich wurden?
Welche Rolle spielten z.B. folgende Gesichtspunkte:
 - Wurden Evaluationsberichte rechtzeitig vorgelegt?
 - Wurden Ergebnisse so präsentiert, dass Sie bzw. Ihre Organisation damit gut umgehen konnten?
 - Haben bisherige Evaluationen alles in allem das geklärt, was aus Sicht Ihrer Organisation geklärt werden sollte?
 - Inwiefern wurden Daten, Schlussfolgerungen und Empfehlungen als gesichertes Wissen erachtet, und wie wurde diese Einschätzung gebildet?
 - Inwiefern standen die Evaluationen mit einem konkreten Orientierungs- oder Entscheidungsbedarf in Zusammenhang?
 - Inwiefern entsprachen Ergebnisse bereits vorhandenen Sichtweisen, und inwiefern waren sie überraschend?
3. Welche Ressourcen, Prozesse und Vorkehrungen gab und gibt es in Ihrer Organisation, um auf die Ergebnisse von durchgeführten Evaluationen zu reagieren und neues Wissen aus den Evaluationen zu verarbeiten?
4. Welcher Stellenwert kam Evaluationen bislang im Verhältnis zu anderen Wissensquellen zu, die Ihrer Organisation in Bezug auf FTI-politische Fragestellungen zur Verfügung stehen?
5. Wie und inwiefern werden Evaluationsergebnisse an andere Akteure des FTI-politischen Systems und Akteure in der FTI-Landschaft weitergegeben? Wie zirkulieren Evaluationsergebnisse im FTI-politischen Governance-System? Inwiefern liegt ein FTI-politisches Informationssystem vor, in dem Evaluationen kontinuierlich zu einem systematischen Wissenszuwachs über das FTI-System beitragen?
6. Wie entstehen konkrete Evaluationsdesigns in Ihrer Organisation? Welche Rolle spielten und spielen dabei rechtliche Rahmen (Gesetze, Verordnungen, Richtlinien), Abstimmungsprozesse intern und extern, und EvaluatorInnen?
7. Welche Möglichkeiten hatten und haben beauftragte Evaluationen während ihres Durchführungszeitraums, um auf veränderte Umstände zu reagieren (z.B. unerwartete Zwischenergebnisse, Veränderungen des untersuchten Programms, Veränderungen der Informationsbedürfnisse des Auftraggebers)?
8. Was wären aus Ihrer Sicht Verbesserungsmöglichkeiten, damit sich der Nutzen von Evaluationen in Zukunft weiter erhöhen kann?

ANHANG 8: Ergebnisse der Online-Umfrage unter FTI-EvaluatorInnen

Die Ergebnisse werden in der Struktur des Fragebogens mit Fragestellungen und Antwortoptionen im Originalwortlaut wiedergegeben.

1. An wievielen Programmevaluationen im Bereich der österreichischen Forschungs-, Technologie- und Innovationspolitik, die von Bundesministerien oder Bundesagenturen beauftragt wurden und seit Anfang 2003 fertig gestellt wurden, waren Sie bislang insgesamt als externer Evaluator/ externe Evaluatorin beteiligt?		
n = 68		
	% aller Antworten	% gültige Antworten
0	27,9	---
1	19,1	26,5
2-3	20,6	28,6
4-5	10,3	14,3
6 oder mehr	22,1	30,6

2. Wenn Sie auf die Programmevaluationen zurück blicken, an denen Sie als externe Evaluatorin mitgewirkt haben: Wie häufig haben diese Evaluationen Ihrem Informationsstand nach die folgenden Formen von Nutzen erzeugt?										
n = 37										
	% aller Antworten					% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie	weiss nicht	Immer	häufig	selten	nie	
Die Evaluationsergebnisse und/oder Empfehlungen wurden vom Auftraggeber zur Entscheidungsfindung über das evaluierte Programm genutzt (Entscheidung über Fortsetzung, Mittelausstattung, Gestaltungsweise, oder Aspekte der Programmdurchführung)	18,9	43,2	13,5	0,0	24,3	25,0	57,1	17,9	0,0	1,93
Die Evaluationsergebnisse haben den Auftraggeberinnen, Entscheidungsträgerinnen über die untersuchten Programme und/oder weiteren Programmbeteiligten geholfen, über den Evaluationsgegenstand zu lernen und neue Sichtweisen darauf zu entwickeln	35,1	35,1	13,5	5,4	10,8	39,4	39,4	15,2	6,1	1,88
Das Vorliegen des Evaluationsberichts bzw. die Tatsache, dass überhaupt evaluiert wurde, diente zur Rechtfertigung bereits zuvor getroffener Entscheidungen.	5,4	35,1	32,4	0,0	27,0	7,4	48,1	44,4	0,0	2,37
Das Vorliegen des Evaluationsberichts bzw. die Tatsache, dass überhaupt evaluiert wurde, diente zur formalen Untermauerung, dass mit dem Programm grundsätzlich rational umgegangen wird.	18,9	43,2	10,8	2,7	24,3	25,0	57,1	14,3	3,6	1,96
Das Vorliegen des Evaluationsberichts bzw. die Tatsache, dass überhaupt evaluiert wurde, diente dazu, andere Akteure in der politischen Sphäre vom Programm zu überzeugen.	16,2	43,2	16,2	0,0	24,3	21,4	57,1	21,4	0,0	2,00
Die Evaluationsergebnisse haben das verfügbare Wissen angereichert, das Akteure im Umfeld des evaluierten Programms nutzen konnten bzw. können (z.B. andere Abteilungen desselben Ministeriums oder derselben Agentur, andere mit FTI befasste Ministerien)	16,2	32,4	29,7	2,7	18,9	20,0	40,0	36,7	3,3	2,23
Das durch die Evaluation erzeugte Wissen kam auch Personen bzw. gesellschaftlichen Sphären zu Gute, die nicht unmittelbar am Programm beteiligt waren oder sind.	5,4	21,6	29,7	8,1	35,1	8,3	33,3	45,8	12,5	2,63
Kognitive, verhaltensförmige oder organisatorische Veränderungen traten bei Auftraggeberinnen und/oder anderen in die Evaluation Einbezogenen bereits im Lauf des Evaluationsprozesses ein, bereits vor Vorliegen der Ergebnisse bzw. unabhängig davon.	8,1	21,6	35,1	8,1	27,0	11,1	29,6	48,1	11,1	2,59

Die Evaluationsergebnisse wurden von den Auftraggeber-Organisationen routinemäßig verarbeitet, ohne dass einer der genannten Effekte zu beobachten gewesen wäre.	0,0	13,5	29,7	24,3	32,4	0,0	20,0	44,0	36,0	3,16
--	-----	------	------	------	------	-----	------	------	------	------

3. Und inwiefern finden Sie, dass folgende Effekte durch die FTI-Programmevaluationen ausgelöst wurden, an denen Sie aktiv mitgewirkt haben?

	% aller Antworten					% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie	weiss nicht	Immer	häufig	selten	nie	
An Evidenz und Verbesserungen orientiertes Denken hat Einzug in die Organisationskultur der auftrag-gebenden und/oder Programm- durchführenden Institutionen gehalten. Evaluation wird mit weniger Skepsis begegnet, man teilt Bedeutungen und Prioritäten.	9,4	37,5	21,9	6,3	25,0	12,5	50,0	29,2	8,3	2,33
Die Interventionen der evaluierten Programme wurden unterstützt oder verstärkt. Es wurden z.B. Outcomes und/oder Impacts der Programme verbessert, oder es wurde Evaluation ins Programm integriert, sodass Reflexivität Teil der Programmpraxis wurde.	6,3	50,0	18,8	3,1	18,8	8,0	64,0	24,0	4,0	2,24
Ein gemeinsames Verständnis innerhalb der evaluierten Programme hat sich vertieft. Man zieht in verantwortlichen und durchführenden Organisationen am selben Strang, und/oder es kam verstärkt zur Ausrichtung von Ressourcen auf Programmprioritäten.	9,4	46,9	12,5	3,1	25,0	13,0	65,2	17,4	4,3	2,13
Aktivitäten, die in den Evaluationen oder in damit zusammenhängenden Monitoringsystemen gemessen wurden, wurden bzw. werden in den Programmen verstärkt gesetzt.	6,3	31,3	25,0	0,0	34,4	10,0	50,0	40,0	0,0	2,30
Es kam es zu einem erhöhten Engagement der Programmbeteiligten, und/oder zu einem verstärkten Sinn für die Eigentümerschaft am Programm unter ihnen.	9,4	34,4	25,0	3,1	25,0	13,0	47,8	34,8	4,3	2,30
Weiterentwicklung der auftraggebenden und/oder Programm- durchführenden Organisation(en), indem Evaluationskapazität aufgebaut wurde, die Anpassungsfähigkeit der Organisation erhöht wurde, oder die Effektivität des Programmanagements gestärkt wurde.	3,1	37,5	25,0	0,0	31,3	4,8	57,1	38,1	0,0	2,33

4. Was war aus Ihrer Sicht für positive Effekte des Evaluierens letztlich mehr ausschlaggebend: bestimmte einzelne Programmevaluationen oder die mittlerweile mehrjährige Evaluationspraxis der betreffenden Organisationen? Bitte wählen Sie die Antwortoption, die aus Ihrer Sicht am Besten zutrifft:

	% aller Antworten	% gültige Antworten
Einzelne Programmevaluationen waren definitiv ausschlaggebend und haben als solche ausgereicht, um Effekte auszulösen.	10,0	14,3
Einzelne Programmevaluationen waren tendenziell ausschlaggebend.	23,3	33,3
Die Tatsache, dass über mehrere Jahre hinweg immer wieder Unterschiedliches evaluiert wurde, war tendenziell stärker verantwortlich.	26,7	38,1
Die Tatsache, dass über mehrere Jahre hinweg immer wieder Unterschiedliches evaluiert wurde, war definitiv ausschlaggebend.	10,0	14,3
weiss nicht / beobachte keine positiven Effekte	30,0	---

5. Welchen Einfluss haben die folgenden Faktoren aus Ihrer Sicht darauf ausgeübt, ob und wie sehr die Programmevaluationen, an denen Sie mitgewirkt haben, genutzt wurden?

n = 32										
	% aller Antworten					% gültige Antworten				Mittelwert
	sehr einflussreich	eher einflussreich	eher nicht einflussreich	gar nicht einflussreich	Trifft nicht zu / weiss nicht	sehr efr.	eher efr.	eher nicht efr.	gar nicht efr.	
Genauigkeit in der Methodenanwendung und Datenanalyse	12,5	40,6	25,0	9,4	12,5	14,3	46,4	28,6	10,7	2,36
Art des Evaluationsansatzes	25,0	40,6	25,0	0,0	9,4	27,6	44,8	27,6	0,0	2,00
Glaubwürdigkeit der EvaluatorenInnen gegenüber den AuftraggeberInnen	59,4	28,1	6,3	0,0	6,3	63,3	30,0	6,7	0,0	1,43
Glaubwürdigkeit der EvaluatorenInnen gegenüber anderen einbezogenen Stakeholdern (z.B. programmdurchführende Agentur) und/oder Zielgruppen des untersuchten Programms	40,6	43,8	6,3	0,0	9,4	44,8	48,3	6,9	0,0	1,62
Objektivität	34,4	25,0	21,9	0,0	18,8	42,3	30,8	26,9	0,0	1,85
Angemessenheit der Evaluationskriterien	25,0	28,1	18,8	3,1	25,0	33,3	37,5	25,0	4,2	2,00
Einsatz von in FTI-Evaluationen gebräuchlichen Methoden	6,3	37,5	34,4	3,1	18,8	7,7	46,2	42,3	3,8	2,42
Methodenvielfalt	15,6	37,5	18,8	6,3	21,9	20,0	48,0	24,0	8,0	2,20
Finesse der Methodenanwendung	12,5	15,6	21,9	18,8	31,3	18,2	22,7	31,8	27,3	2,68
Beleuchtung bestimmter Fakten durch mehrere parallel eingesetzte Methoden	15,6	31,3	25,0	6,3	21,9	20,0	40,0	32,0	8,0	2,28
Erstellung von Logic Charts oder anderen Darstellungen der Programmlogik	6,3	37,5	25,0	6,3	25,0	8,3	50,0	33,3	8,3	2,42
Klarheit der Berichterstattung an die HauptadressatInnen der Evaluation	53,1	34,4	3,1	0,0	9,4	58,6	37,9	3,4	0,0	1,45
Präsentationsweise, Stil der Berichterstattung	34,4	43,8	6,3	3,1	12,5	39,3	50,0	7,1	3,6	1,75
Breite der Dissemination von Ergebnissen	0,0	25,0	43,8	9,4	21,9	0,0	32,0	56,0	12,0	2,80
Ausgewogene Darstellung von Stärken und Schwächen des untersuchten Programms	18,8	50,0	12,5	3,1	15,6	22,2	59,3	14,8	3,7	2,00
Zeitgerechte Übermittlung der Evaluationsergebnisse gemäß vertraglichen Vereinbarungen	15,6	25,0	34,4	9,4	15,6	18,5	29,6	40,7	11,1	2,44
Zeitgerechte Übermittlung der Evaluationsergebnisse im Verhältnis zu konkreten Entscheidungsfindungsprozessen	28,1	34,4	9,4	3,1	25,0	37,5	45,8	12,5	4,2	1,83
Planungen hinsichtlich der Nutzung in einer Frühphase der Evaluation	15,6	25,0	21,9	6,3	31,3	22,7	36,4	31,8	9,1	2,27
Zuschnitt auf Informationsbedürfnisse der Entscheidungsträger über das untersuchte Programm	43,8	34,4	6,3	3,1	12,5	50,0	39,3	7,1	3,6	1,64
Klare Darstellung der Prinzipien, warum welche Daten und Verfahren zur Beantwortung welcher Fragestellungen herangezogen wurden	15,6	37,5	21,9	12,5	12,5	17,9	42,9	25,0	14,3	2,36
Durchführung einer möglichst umfassenden Analyse, auch wenn im Rahmen der gegebenen Ressourcen nicht für alle behandelten Aspekte Daten vorlagen bzw. erhoben werden konnten	3,1	34,4	28,1	12,5	21,9	4,0	44,0	36,0	16,0	2,64
Beschränkung der Analyse auf Aspekte, für die ausreichend gesicherte Daten vorlagen bzw. im Rahmen der gegebenen Ressourcen erhoben werden konnten	12,5	25,0	34,4	6,3	21,9	16,0	32,0	44,0	8,0	2,44

6. Und welchen Einfluss haben die folgenden Kontextfaktoren aus Ihrer Sicht darauf ausgeübt, ob und wie sehr die Ergebnisse von Programmevaluationen, an denen Sie mitgewirkt haben, genutzt wurden?

n = 32

	% aller Antworten					% gültige Antworten				Mittelwert
	sehr einflussreich	eher einflussreich	eher nicht einflussreich	gar nicht einflussreich	Trifft nicht zu / weiss nicht	sehr efr.	eher efr.	eher nicht efr.	gar nicht efr.	
Erwartung der Auftraggeberinnen, dass die Evaluation ihnen und ihren Vorhaben von Nutzen sein wird	56,3	31,3	0,0	3,1	9,4	62,1	34,5	0,0	3,4	1,45
Erwartung von Stakeholdern des Programms, dass die Evaluation ihnen und ihren Vorhaben von Nutzen sein wird	31,3	40,6	12,5	3,1	12,5	35,7	46,4	14,3	3,6	1,86
Konsistenz der Evaluationsergebnisse mit Sichtweisen und Erwartungen seitens der Auftraggeberinnen	28,1	43,8	9,4	3,1	15,6	33,3	51,9	11,1	3,7	1,85
Gleichzeitiges Vorhandensein anderer, unter Umständen auch konfligierender Informationsquellen zu den FTI-politischen Fragestellungen bei den Auftraggeber-Organisationen	15,6	37,5	18,8	3,1	25,0	20,8	50,0	25,0	4,2	2,13
Direkter Konnex der Evaluationen mit einem aktuellen Entscheidungsbedarf oder Problemdruck	43,8	31,3	12,5	0,0	12,5	50,0	35,7	14,3	0,0	1,64
Begleitung der Evaluationen durch Evaluationsmanagerinnen in den Auftraggeber-Organisationen	3,1	31,3	31,3	6,3	28,1	4,3	43,5	43,5	8,7	2,57
Bandbreite der verschiedenen Adressatinnen bzw. Adressatengruppen mit unterschiedlichen Informationsbedürfnissen	3,1	25,0	28,1	21,9	21,9	4,0	32,0	36,0	28,0	2,88
Wichtigkeit bzw. Tragweite der mit der Evaluation verbundenen Entscheidung	31,3	31,3	12,5	9,4	15,6	37,0	37,0	14,8	11,1	2,00
Rolle der direkten Auftraggeberin als Person in ihrer Organisation	21,9	40,6	3,1	12,5	21,9	28,0	52,0	4,0	16,0	2,08
Erfahrung der Auftraggeberin als Person mit Evaluation	21,9	50,0	18,8	0,0	9,4	24,1	55,2	20,7	0,0	1,97
Erfahrung der Organisation mit Evaluation	25,0	50,0	12,5	3,1	9,4	27,6	55,2	13,8	3,4	1,93
Persönliche Sichtweisen und Denkstile der direkten Auftraggeberin	31,3	40,6	12,5	0,0	15,6	37,0	48,1	14,8	0,0	1,78
Ressourcen und organisatorische Anpassungen für die Verarbeitung von Evaluationsergebnissen in den Auftraggeber- Organisationen und/oder in die Evaluationen einbezogenen Organisationen	28,1	31,3	9,4	6,3	25,0	37,5	41,7	12,5	8,3	1,92
Reifegrad des Programms hinsichtlich Evaluierbarkeit	18,8	34,4	28,1	6,3	12,5	21,4	39,3	32,1	7,1	2,25
Zirkulieren von Evaluationsergebnissen zwischen der Auftraggeber-Organisation und anderen Organisationen im FTI- politischen Governance-System	6,3	34,4	31,3	3,1	25,0	8,3	45,8	41,7	4,2	2,42
Informelle Rollen von Personen und/oder Netzwerken bei der Verbreitung und Verarbeitung von Evaluationsergebnissen	12,5	21,9	34,4	9,4	21,9	16,0	28,0	44,0	12,0	2,52
Rivalitäten innerhalb der auftraggebenden Organisation, zwischen dieser und anderen beteiligten Organisationen, oder zwischen beteiligten Organisationen	12,5	21,9	31,3	0,0	34,4	19,0	33,3	47,6	0,0	2,29
Abhängigkeit der Entscheidungsträgerinnen über das Programm von Geldgebern	9,4	28,1	28,1	0,0	34,4	14,3	42,9	42,9	0,0	2,29
Anzweiflung der Ergebnisse durch wichtige Stakeholder der untersuchten Programme auf Grund ihrer vorhandenen Sichtweisen	3,1	28,1	28,1	15,6	25,0	4,2	37,5	37,5	20,8	2,75
Grundsätzlich geringe Neigung von Entscheidungsträgerinnen, sich auf Evaluationsergebnisse zu stützen	9,4	12,5	40,6	12,5	25,0	12,5	16,7	54,2	16,7	2,75

7. Möchten Sie etwas zu den Nutzungsweisen der Programmevaluationen oder zu Einflüssen auf die Nutzung ergänzen oder genauer ausführen? Beobachten Sie vielleicht einen Trend? Bitte nutzen Sie den folgenden Platz für Ihre Angaben!

n = 49		
	% aller Antworten	% inhaltliche Antworten
keine Angabe	79,6	----
Inhaltlich verwertbare Angaben	14,4	70,0
Angaben, die sich nicht in verwertbarer Weise auf die Fragestellung beziehen	0,6	30,0

8. Wie häufig gab es bei den Programmevaluationen, an denen Sie mitgewirkt haben, folgende Schritte, um die Auftraggeberinnen/ Auftraggeber zu ermuntern, sich für die Evaluation und ihre Ergebnisse zu interessieren?

	% aller Antworten					% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie	Trifft nicht zu / weiss nicht	Immer	häufig	selten	nie	
Es wurden Zwischenberichte bereitgestellt	43,8	37,5	15,6	3,1	0,0	43,8	37,5	15,6	3,1	1,78
Schriftliche Berichte wurden durch mündliche Kommunikation ergänzt und unterstützt	71,9	21,9	3,1	3,1	0,0	71,9	21,9	3,1	3,1	1,38
Die Adressatinnen wurden während der Laufzeit der Evaluation kontinuierlich informiert	37,5	37,5	18,8	6,3	0,0	37,5	37,5	18,8	6,3	1,94
Mögliche Nutzungen der Ergebnisse wurden antizipiert, gegebenenfalls priorisiert, und gezielt bedient	15,6	43,8	28,1	12,5	0,0	15,6	43,8	28,1	12,5	2,38
Es fanden Abstimmungen zu den eingesetzten Methoden statt (z.B. Befragungsbasis, Interviewpartner, ...)	40,6	40,6	12,5	3,1	3,1	41,9	41,9	12,9	3,2	1,77
Es fanden während der Laufzeit der Evaluation Gespräche über veränderte Informationsbedürfnisse statt	25,0	37,5	25,0	9,4	3,1	25,8	38,7	25,8	9,7	2,19
Es fanden Feedback-Gespräche statt, um sich mit Zwischenergebnissen bzw. dem Fortschritt der Evaluation auseinander zu setzen	34,4	46,9	9,4	6,3	3,1	35,5	48,4	9,7	6,5	1,87
Es wurde Gelegenheit zur Stellungnahme zu Schlussfolgerungen und Empfehlungen gegeben	65,6	18,8	6,3	3,1	6,3	70,0	20,0	6,7	3,3	1,43
Es gab Arrangements für eine Nachbereitung, um Unterstützung bei der Interpretation und Anwendung der Ergebnisse und Empfehlungen zu geben	34,4	28,1	21,9	6,3	9,4	37,9	31,0	24,1	6,9	2,00

9. Und wie häufig wurden bei den Programmevaluationen, an denen Sie mitgewirkt haben, folgende Schritte gesetzt, um weitere, neben den Auftraggeberinnen einbezogene Beteiligte zu ermuntern, sich für die Evaluation und ihre Ergebnisse zu interessieren?

n = 30											
	% aller Antworten						% gültige Antworten				Mittelwert
					Trifft nicht zu / weiss nicht						
	Immer	häufig	selten	nie		Immer	häufig	selten	nie		
Die Evaluationszwecke wurden an sie kommuniziert, um das Einbringen von Sichtweisen und Standpunkten zu ermöglichen	20,0	26,7	33,3	13,3	6,7	21,4	28,6	35,7	14,3		2,43
Es wurden Zwischenberichte übermittelt	13,3	23,3	26,7	30,0	6,7	14,3	25,0	28,6	32,1		2,79

Schriftliche Berichte wurden durch mündliche Kommunikation ergänzt und unterstützt	13,3	20,0	43,3	16,7	6,7	14,3	21,4	46,4	17,9	2,68
Die Adressatinnen wurden während der Laufzeit der Evaluation kontinuierlich informiert	3,3	20,0	43,3	23,3	10,0	3,7	22,2	48,1	25,9	2,96
Es fanden Feedback-Gespräche statt, um sich mit Zwischenergebnissen bzw. dem Fortschritt der Evaluation auseinander zu setzen	3,3	23,3	40,0	23,3	10,0	3,7	25,9	44,4	25,9	2,93
Es fanden während der Laufzeit der Evaluation Gespräche über veränderte Informationsbedürfnisse statt	6,7	16,7	36,7	30,0	10,0	7,4	18,5	40,7	33,3	3,00
Es wurde Gelegenheit zur Stellungnahme zu Schlussfolgerungen und Empfehlungen gegeben	10,0	33,3	33,3	20,0	3,3	10,3	34,5	34,5	20,7	2,66
Direkte Übermittlung des Endberichts oder eigens zugeschnittener Kurzinformationen	16,7	36,7	23,3	10,0	10,0	19,2	42,3	26,9	11,5	2,31
Veranstaltung von gezielten Präsentationen, Diskussionsrunden, Workshops	16,7	36,7	26,7	16,7	3,3	17,2	37,9	27,6	17,2	2,45
Arrangements für eine Nachbereitung, um Unterstützung bei der Interpretation und Anwendung der Ergebnisse und Empfehlungen zu geben	10,0	20,0	33,3	26,7	6,7	11,1	22,2	37,0	29,6	2,85

10. Wie häufig wurden folgende Gruppen in der Planungsphase der Evaluation einbezogen, um ihre Informationsbedürfnisse und Sichtweisen auf den Evaluationsgegenstand festzustellen?											
n = 28											
	% aller Antworten					Trifft nicht zu / weiss nicht	% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie	Immer		häufig	selten	nie		
Für die Evaluation zuständige Ansprechpartnerinnen in der Auftraggeber- Organisation	64,3	21,4	7,1	0,0	7,1	69,2	23,1	7,7	0,0	1,38	
Personen oder Personengruppen, die abgesehen von der direkten Auftraggeberin oder neben ihr für die Konzeption und Gestaltung des Programms verantwortlich waren bzw. über die Zukunft des untersuchten Programms zu entscheiden hatten	35,7	39,3	10,7	7,1	7,1	38,5	42,3	11,5	7,7	1,88	
Personen oder Personengruppen, die mit der Umsetzung des Programms befasst waren	39,3	50,0	3,6	3,6	3,6	40,7	51,9	3,7	3,7	1,70	
Personen oder Personengruppen, die durch das Programm erreicht werden sollten	21,4	10,7	28,6	32,1	7,1	23,1	11,5	30,8	34,6	2,77	
Personen oder Personengruppen, die durch das Programm oder durch Veränderungen des Programms Nachteile erleiden hätten können	7,1	7,1	21,4	28,6	32,1	11,1	11,1	33,3	44,4	3,11	
Personen oder Personengruppen, die ähnliche Programme planten bzw. in ihren FTI-politischen Rollen Interesse am untersuchten Programm hatten bzw. haben konnten	7,1	3,6	39,3	25,0	21,4	9,5	4,8	52,4	33,3	3,10	

11. Im Folgenden finden Sie einige Aussagen zur Gestaltung von Evaluationen. Bitte geben Sie jeweils an, wie sehr die Aussage im Gesamtbild auf die Programmevaluationen zutrifft, an denen Sie aktiv mitgewirkt haben:

n = 28

	% aller Antworten					Trifft nicht zu / weiss nicht	% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie			Immer	häufig	selten	nie	
Es war klar verankert, welche Zwecke mit der Evaluation verfolgt wurden, sodass das Evaluationsteam einen klaren Arbeitsauftrag hatte.	39,3	50,0	10,7	0,0	0,0		39,3	50,0	10,7	0,0	1,71
Es stand ein Hauptzweck der Evaluation klar im Vordergrund bzw. die Evaluationszwecke wurden gemeinsam mit den Auftraggeberinnen mit klaren Prioritäten ausgestattet.	28,6	60,7	10,7	0,0	0,0		28,6	60,7	10,7	0,0	1,82
Es wurde bei der Evaluationsplanung explizit reflektiert, wie die Ergebnisse später genutzt werden sollten und wie die Evaluationsergebnisse diesbezüglich verbreitet werden sollten.	17,9	28,6	46,4	3,6	3,6		18,5	29,6	48,1	3,7	2,37
Es wurde bei der Evaluationsplanung explizit reflektiert, wie die Evaluation Lernen oder organisatorischen Wandel unterstützen sollte, und darauf ausgerichtete Schritte wurden vorgesehen.	17,9	17,9	50,0	10,7	3,6		18,5	18,5	51,9	11,1	2,56
Es wurden zum selben Programm auch noch andere Evaluationen mit unterschiedlicher Zwecksetzung durchgeführt, gegebenenfalls durch andere Evaluationsteams.	3,6	7,1	25,0	53,6	10,7		4,0	8,0	28,0	60,0	3,44
Der Evaluationsprozess verfügte über eine Eingangsphase, in der Fragestellungen, Datenlagen und Methoden reflektiert wurden, um Entscheidungen über das endgültige Evaluationsdesign gemeinsam mit den Auftraggeberinnen zu treffen.	35,7	32,1	25,0	3,6	3,6		37,0	33,3	25,9	3,7	1,96
Die Evaluation verfügte über eingeplante Flexibilität, um während ihres Durchführungszeitraums auf veränderte Umstände zu reagieren (z.B. unerwartete Datenlage, unerwartete Zwischenergebnisse, Veränderungen des untersuchten Programms, etc.).	17,9	42,9	28,6	10,7	0,0		17,9	42,9	28,6	10,7	2,32
Der Endbericht wurde in vollständiger Fassung einschließlich aller Anhänge ohne irgendeine Abänderung publiziert.	14,3	50,0	25,0	10,7	0,0		14,3	50,0	25,0	10,7	2,32
Ein auftraggeberseitig eigens produzierter Bericht wurde statt des Original- Endberichts publiziert.	3,6	7,1	25,0	57,1	7,1		3,8	7,7	26,9	61,5	3,46
Es wurde bei der Evaluationsplanung explizit vorgesehen, auch Erkenntnisse bereitzustellen, die der FTI-politischen, wissenschaftlichen oder öffentlichen Diskussion Anregung geben können.	14,3	25,0	39,3	14,3	7,1		15,4	26,9	42,3	15,4	2,58
Es wurden mehrere Berichte und/oder Präsentationen unterschiedlichen Zuschnitts für verschiedene Adressatinnen bzw. Adressatengruppen erstellt.	10,7	7,1	42,9	32,1	7,1		11,5	7,7	46,2	34,6	3,04
Es wurden Diskussionsveranstaltungen abgehalten oder in anderer Weise Schritte gesetzt, um auch nicht direkt in die Evaluation eingebundenen Personen bzw. Personengruppen die Möglichkeit zu geben, aufgeklärte Beobachterinnen zu werden.	10,7	21,4	39,3	21,4	7,1		11,5	23,1	42,3	23,1	2,77
Die Möglichkeit, Informationen freizugeben oder zu veröffentlichen, unterlag ausschließlich der Kontrolle der Auftraggeberin.	67,9	7,1	7,1	7,1	10,7		76,0	8,0	8,0	8,0	1,48

12. Im Zuge einer Evaluation sind regelmäßig Fakten in der Ergebnisinterpretation zu bewerten. Bitte geben Sie jeweils an, wie sehr die folgenden Aussagen im Gesamtbild auf die Programmevaluationen zutreffen, an denen Sie aktiv mitgewirkt haben:

n = 25

	% aller Antworten					Trifft nicht zu / weiss	% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie	nicht		Immer	häufig	selten	nie	
Die Evaluationen waren in ihrer Herangehensweise grundsätzlich wertfrei.	32,0	52,0	8,0	4,0	4,0		33,3	54,2	8,3	4,2	1,83
Die Evaluatorinnen waren alleine dafür zuständig, die aufgearbeiteten Fakten nach eigenen Überlegungen zu bewerten.	28,0	52,0	16,0	4,0	0,0		28,0	52,0	16,0	4,0	1,96
Es wurde in einer Frühphase der Evaluation in Abstimmung mit den Auftraggeberinnen festgelegt, welche Wertmaßstäbe später herangezogen werden sollten, um die Ergebnisse zu interpretieren.	4,0	36,0	28,0	28,0	4,0		4,2	37,5	29,2	29,2	2,83
Die Bewertungskriterien, die die Auftraggeberinnen vorgaben bzw. deren Sichtweisen entsprachen, wurden angewendet.	16,0	36,0	20,0	24,0	4,0		16,7	37,5	20,8	25,0	2,54
Eine mögliche Bandbreite an begründeten Wertvorstellungen, die zum Programm unter verschiedenen Akteuren existieren konnte, wurde im Rahmen der Evaluationen untersucht und berichtet.	12,0	44,0	24,0	16,0	4,0		12,5	45,8	25,0	16,7	2,46
Es wurden unterschiedliche Zugangsweisen dazu, wie die Fakten in der Ergebnisinterpretation zu bewerten sein werden, in einer Frühphase der Evaluation gegeneinander abgewogen und zwischen ihnen entschieden.	12,0	24,0	40,0	24,0	0,0		12,0	24,0	40,0	24,0	2,76
Bewertungsmaßstäbe, die bei den Auftraggeberinnen oder bei anderen Stakeholdern des Programms möglicherweise existierten, waren für die konkrete Vorgehensweise nicht wesentlich.	16,0	24,0	40,0	12,0	8,0		17,4	26,1	43,5	13,0	2,52
Die durch die Evaluation erbrachten Fakten wurden nicht alleine durch die Evaluatorinnen bewertet, sondern in einem speziellen Verfahren (z.B. gemeinsam mit Auftraggeberinnen, gemeinsam mit Stakeholdern, durch eine Steuerungsgruppe oder Jury).	4,0	28,0	28,0	36,0	0,0		4,2	29,2	29,2	37,5	3,00

13. Und inwiefern treffen folgende Aussagen im Gesamtbild auf die Programmevaluationen zu, an denen Sie aktiv mitgewirkt haben?

n = 27

	% aller Antworten					Trifft nicht zu / weiss nicht	% gültige Antworten				Mittelwert
	Immer	häufig	selten	nie			Immer	häufig	selten	nie	
Die Auftraggeber-seitigen Erwartungen an die Evaluationen waren unter Berücksichtigung der Ressourcen und Methoden von Anfang an realistisch.	11,1	44,4	40,7	0,0	3,7		11,5	46,2	42,3	0,0	2,31
In der Evaluationsplanung war es möglich, Auswahl und Umfang der von der Evaluation erfassten Informationen so zu bestimmen, dass alle vorgesehenen Fragestellungen gut behandelt werden konnten.	22,2	48,1	18,5	7,4	3,7		23,1	50,0	19,2	7,7	2,12
Das Evaluationsdesign beruhte auf Methoden, die die Auftraggeberin eingesetzt sehen wollte.	11,1	48,1	29,6	7,4	3,7		11,5	50,0	30,8	7,7	2,35
Das Evaluationsdesign beruhte auf Methoden, die ich bzw. mein Institut regelmäßig einsetzen	18,5	70,4	3,7	3,7	3,7		19,2	73,1	3,8	3,8	1,92
Das Evaluationsdesign konnte darauf abgestimmt werden, dass alle Fragestellungen gut behandelt werden konnten, die das Evaluationsteam für wichtig erachtete, auch wenn manche davon ursprünglich nicht von der Auftraggeberin vorgesehen waren.	25,9	37,0	22,2	7,4	7,4		28,0	40,0	24,0	8,0	2,12
Bei der Evaluationsplanung wurde herausgearbeitet, was die wesentlichsten Evaluationsfragestellungen sein sollten, welche Informationen zur deren Beantwortung unabdingbar waren, und was weniger Priorität haben sollte.	44,4	37,0	11,1	3,7	3,7		46,2	38,5	11,5	3,8	1,73
Es wurden Schritte unternommen, um Übereinstimmungen und Abweichungen bezüglich der Evaluationszwecke und -verfahren zwischen Evaluationsteam und Auftraggeberinnen festzustellen und abzuschätzen.	25,9	29,6	25,9	7,4	11,1		29,2	33,3	29,2	8,3	2,17
Ein Logic Model oder andere Darstellung der Wirklogik des Programms wurde erarbeitet, um zur Feindefinition des Evaluationsdesigns und zur Auswahl der benötigten Informationen zu dienen.	7,4	33,3	37,0	18,5	3,7		7,7	34,6	38,5	19,2	2,69
Als aufgrund zeitlicher oder finanzieller Beschränkungen einige für die Evaluation relevante Daten nicht zugänglich wurden, wurde darüber berichtet, einschließlich einer Einschätzung der Auswirkungen auf die Gesamtanalyse des Programms.	11,1	33,3	22,2	14,8	18,5		13,6	40,9	27,3	18,2	2,50
Als im Zuge des Evaluationsprozesses die Vorgehensweisen angepasst und abgeändert wurden, wurde dies im Endbericht dargestellt.	22,2	22,2	25,9	7,4	22,2		28,6	28,6	33,3	9,5	2,24
Bei der Beurteilung von Stärken und Schwächen des Programms wurden auch andere begründete Sichtweisen in Betracht gezogen, die möglicher Weise zu anderen Schlussfolgerungen geführt hätten, und die betreffenden Abwägungen wurden im Bericht dargestellt.	25,9	29,6	18,5	11,1	14,8		30,4	34,8	21,7	13,0	2,17
Um die Fragestellungen der Evaluation herauszufiltern bzw. zu gewichten, wurden verfügbare Materialien wie Forschungsberichte, theoretische Abhandlungen, oder frühere Evaluationen herangezogen.	51,9	40,7	0,0	0,0	7,4		56,0	44,0	0,0	0,0	1,44
Alle für das akkordierte Evaluationsdesign benötigten Daten wurden von den Auftraggebern ohne Vorbehalte und Einschränkungen herausgegeben, soweit verfügbar.	40,7	37,0	7,4	3,7	11,1		45,8	41,7	8,3	4,2	1,71
Wesentliche Daten, die bei anderen Institutionen und Behörden vorlagen, wurden durch geeignete Schritte der Auftraggeber zugänglich gemacht.	22,2	40,7	14,8	3,7	18,5		27,3	50,0	18,2	4,5	2,00
Die Möglichkeiten für ein Evaluationsdesign, das das	11,1	33,3	25,9	14,8	14,8		13,0	39,1	30,4	17,4	2,52

Evaluationsteam als am Besten geeignet erachtet hätte, waren durch die für die Evaluation verfügbaren Ressourcen deutlich eingeschränkt.										
Veränderungen in den Informationsbedürfnissen von Auftraggebern, die während der Laufzeit der Evaluation eintraten, sollten ohne Reflexion bzw. Anpassung der eingesetzten Ressourcen und Verfahren aufgegriffen werden.	3,7	33,3	14,8	25,9	22,2	4,8	42,9	19,0	33,3	2,81
Die Evaluationen waren so angelegt, dass sie sich auch mit nicht-intendierten Wirkungen oder unerwünschten Nebenwirkungen des untersuchten Programms auseinander setzen konnten.	29,6	29,6	18,5	14,8	7,4	32,0	32,0	20,0	16,0	2,20
Als mehrdeutige Ergebnisse vorlagen und/oder mehrere alternative Sichtweisen auf die Bedeutung der Ergebnisse möglich waren, wurde darüber berichtet.	40,7	40,7	7,4	0,0	11,1	45,8	45,8	8,3	0,0	1,63
Bei der Evaluationsdurchführung haben die Evaluatorinnen auch die Rolle einer Vermittlerin, Beraterin, Moderatorin, oder Prozessbegleiterin eingenommen.	14,8	55,6	14,8	7,4	7,4	16,0	60,0	16,0	8,0	2,16

14. Was wären aus Ihrer Sicht Verbesserungsmöglichkeiten, damit sich der Nutzen von Programmevaluationen im österreichischen FTI-Bereich in Zukunft weiter erhöhen kann? Bitte um Ihre Angaben!

n = 49

	% aller Antworten	% inhaltliche Antworten
keine Angabe	79,5	----
Inhaltlich verwertbare Angaben	12,3	60,0
Angaben, die sich nicht in verwertbarer Weise auf die Fragestellung beziehen	8,2	40,0

15. Seit wann führen Sie Evaluationen durch (ungeachtet des Typs oder des Politikfelds)?

n = 27

	% aller Antworten	% gültige Antworten
1-3 Jahre	0,0	0,0
4-6 Jahre	14,8	14,8
7 Jahre oder mehr	85,2	85,2

16. Welche der folgenden Aussagen trifft auf Sie zu? (Mehrfachantworten möglich)

n = 27

	% aller Antworten	% gültige Antworten
Evaluation, Analyse und Politikberatung im FTI-Bereich sind meine Haupttätigkeit	51,9	51,9
Evaluation ist meine Haupttätigkeit und erstreckt sich auch auf andere Politikfelder als FTI	29,6	29,6
Ich habe in meiner Organisation eine leitende Rolle bei der Planung und Durchführung von FTI-Evaluationen inne	51,9	51,9

17. Würden Sie sagen, dass es für die FTI-Programmevaluationen, an denen Sie mitgewirkt haben, eine Leitwissenschaft bzw. ein Leitparadigma gab, durch die bzw. das die Analysestrategie geprägt war?		
n = 25		
	% aller Antworten	% inhaltliche Antworten
keine Leitwissenschaft/ kein Leitparadigma	64,0	---
Methoden quantitativ/qualitativ	16,0	44,4
Skepsis gegenüber Evaluationspraxis	12,0	33,3
Systemische Orientierung	4,0	11,1
Action Research	4,0	11,1
Politikevaluation	4,0	11,1