

**Forschung
Projekt F 2105**

St. Pennig
J. Vogt

**Entwicklung und Erprobung des
Human Resources Performance Modells
zur ökonomischen Evaluation von
Maßnahmen in den Bereichen
Humanfaktoren (HF),
Humanressourcen (HR) und Training (T)**

Dortmund/Berlin/Dresden 2007

Diese Veröffentlichung ist der Abschlussbericht zum Projekt „Guidelines for the Economic Evaluation of Human Factors, Human Resources and Training Programmes“ im Auftrag der European Organisation for the Safety of Air Navigation (Eurocontrol project AO/80/HQ/DK04) und der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA Projekt F 2105). Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Autoren: Stefan Pennig¹
Joachim Vogt²

¹ context

Am Stadtwald 26, 45219 Essen-Kettwig

² Universität Kopenhagen

Østerfarimagsgade 5A, 1353 Kopenhagen, Dänemark

Herausgeber: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin
Friedrich-Henkel-Weg 1-25, 44149 Dortmund
Telefon: 0231 9071-0
Telefax: 0231 9071-2454
E-Mail: poststelle@baua.bund.de
Internet: www.baua.de

Berlin:

Nöldnerstr. 40-42, 10317 Berlin

Telefon: 030 51548-0

Telefax: 030 51548-4170

Dresden:

Proschhübelstr. 8, 01099 Dresden

Telefon: 0351 5639-50

Telefax: 0351 5639-5210

Alle Rechte einschließlich der fotomechanischen Wiedergabe und des auszugsweisen Nachdrucks vorbehalten.

Inhaltsverzeichnis

| | Seite |
|--|-------|
| Kurzreferat | 6 |
| Abstract | 8 |
| Danksagung | 10 |
| 1 Einleitung | 11 |
| 1.1 Notwendigkeit ökonomischer Evaluation von Personalmaßnahmen | 11 |
| 1.2 Entstehung des Projekts | 13 |
| 1.3 Produkte des Projekts | 13 |
| 1.3.1 Literaturübersicht | 14 |
| 1.3.2 Eigene empirische Untersuchungen | 15 |
| 1.3.3 Handlungshilfe | 16 |
| 1.3.4 <i>Economic Evaluation Selection Tool EMS</i> | 17 |
| 1.3.5 <i>Frequently Asked Questions FAQ und Marketing</i> | 17 |
| 1.4 Das <i>Human Resources Performance</i> Modell HPM | 18 |
| 1.5 Zusammenfassung und Überleitung | 25 |
| 2 Empirische Untersuchungen | 26 |
| 2.1 Gesamtkonzeption der drei empirischen Untersuchungen | 26 |
| 2.2 <i>Critical Incident Stress Management CISM</i> | 27 |
| 2.2.1 Flugsicherung | 27 |
| 2.2.2 Zusammenfassung der Machbarkeitsstudie | 29 |
| 2.2.3 Ziele der Hauptstudie | 31 |
| 2.2.4 Der Evaluationsprozess in der Hauptstudie | 32 |
| 2.2.5 Modellierung der Zusammenhänge von Verhalten und Wirtschaftlichkeit in der Hauptstudie und Konzeption der Erhebungsinstrumente | 34 |
| 2.2.5.1 <i>Workshops</i> mit den Fachkräften | 35 |
| 2.2.5.2 Interviews mit den Führungskräften | 39 |
| 2.2.5.3 Fragebogen für Fluglotsen | 42 |
| 2.2.5.4 Fragebogen für <i>Peers</i> | 47 |
| 2.2.6 Stichprobenbeschreibung und Rücklauf | 49 |
| 2.2.7 Auswertung der Fragebogen | 51 |
| 2.2.7.1 Selbststeuerung der Fluglotsen | 51 |
| 2.2.7.2 Einfluss der operativen Steuerung | 67 |
| 2.2.7.3 Bedeutung für die strategische Steuerung | 73 |
| 2.2.7.4 Langfristige Folgen von kritischen Ereignissen | 77 |
| 2.2.7.5 Wechselwirkungen von CISM und Sicherheitskultur | 80 |
| 2.2.8 Resumée zur CISM-Studie | 85 |

| | | |
|---------|--|-----|
| 2.3 | Programmevaluation: Personalentwicklungsmaßnahmen im Call Center | 86 |
| 2.3.1 | Hintergrund und Zielsetzung | 86 |
| 2.3.2 | Datenmodell für die Programmevaluation | 86 |
| 2.3.3 | Planung und Konzeption der Programmevaluation | 88 |
| 2.3.4 | Untersuchte Stichprobe und Rücklauf | 91 |
| 2.3.5 | Ergebnisse | 91 |
| 2.3.5.1 | Welche Anforderungen sind besonders bedeutsam für die Beurteilung der Arbeit des <i>Call Center</i> Agenten? | 91 |
| 2.3.5.2 | Entsprechen die durch die verschiedenen Maßnahmen entwickelten Kompetenzen den Anforderungen? | 93 |
| 2.3.5.3 | Welche Merkmale des Anforderungsprofils werden durch die einzelnen Maßnahmen abgedeckt? | 97 |
| 2.3.5.4 | Wie effektiv werden durch die Maßnahmen die beabsichtigten Entwicklungen gefördert? | 99 |
| 2.3.5.5 | Welchen Effekt haben die initiierten und unterstützen Kompetenzentwicklungen auf die Tätigkeitsausübung und die Arbeitsergebnisse? | 101 |
| 2.3.5.6 | Welchen Einfluss haben die Kontextfaktoren auf die Qualität der Arbeit? | 106 |
| 2.3.5.7 | Wie sehr können durch Maßnahmen der Kompetenzförderung Ergebniseffekte erwartet werden? | 107 |
| 2.3.6 | Zusammenfassung und Resumée zur Programmevaluation | 109 |
| 2.4 | Kurzevaluation: <i>E-learning</i> | 112 |
| 2.4.1 | Hintergrund und Zielsetzung | 112 |
| 2.4.2 | Untersuchungspersonen und –methoden | 112 |
| 2.4.3 | Ergebnisse | 114 |
| 2.4.3.1 | Fähigkeiten, Arbeitstätigkeiten und Leistung | 115 |
| 2.4.3.2 | <i>E-learning</i> im Vergleich zu Simulator und Klassenunterricht | 118 |
| 2.4.3.3 | Monetarisierung von <i>E-learning</i> als Ersatz für Simulator/Klasse | 120 |
| 2.4.3.4 | <i>Return on Investment</i> und <i>Break Even</i> der <i>E-learning</i> Nutzung | 122 |
| 2.4.3.5 | Perspektiven der Verbesserung von <i>E-learning</i> | 126 |
| 2.4.4 | Zusammenfassung der Ergebnisse | 128 |
| 2.4.5 | Resumée zur Kurzevaluation | 129 |
| 3 | Gesamtdiskussion | 131 |
| 4 | Literaturverzeichnis | 140 |
| 5 | Abbildungs- und Tabellenverzeichnis | 142 |
| 5.1 | Abbildungen | 142 |
| 5.2 | Tabellen | 146 |
| 6 | Formelzeichen und Abkürzungen | 147 |

- Anh. 1: CISM Interviewleitfaden Manager¹
- Anh. 2: CISM Lotsenfragebogen
- Anh. 3: CISM Peerfragebogen
- Anh. 4: PE-Programm Mitarbeiterfragebogen
- Anh. 5: PE-Programm Führungskräftefragebogen
- Anh. 6: *E-learning* Interviewleitfaden Manager
- Anh. 7: *E-learning* Traineefragebogen
- Anh. 8: *E-learning* Trainerfragebogen

¹ Alle Anhänge sind auf Anfrage hin bei den Autoren verfügbar.

Entwicklung und Erprobung des *Human Resources Performance* Modells zur ökonomischen Evaluation von Maßnahmen in den Bereichen Humanfaktoren (HF), Humanressourcen (HR) und Training (T)

Kurzreferat

Dieses Dokument ist der wissenschaftliche Abschlussbericht des Projekts F2105 „*Guidelines for the Economic Evaluation of Human Factors HF, Human Resources HR and Training T Programmes*“, das die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin gemeinsam mit der *European Organisation for the Safety of Air Navigation* EUROCONTROL initiierte.

Innerhalb dieses Projekts wurden Methoden der ökonomischen Evaluation von HF/HR/T-Maßnahmen in einer Literaturübersicht zusammengefasst. Weder in wissenschaftlichen noch in praxisorientierten Quellen konnte ein Ansatz identifiziert werden, der Methoden der Wirtschafts- und der Verhaltenswissenschaften zufrieden stellend integriert und gleichermaßen wissenschaftlichen Anforderungen wie Bedürfnissen der Organisationspraxis gerecht wird. Die ökonomischen und strategischen Wirkungen von HF, HR und T im Unternehmen wurden bisher zu wenig dokumentiert.

Die Stärken der verschiedenen, bisher wenig integrierten Ansätze zum Thema wurden daher in einem neuen und umfassenden Rahmenkonzept zusammengeführt, dem *Human Resources Performance* Modell (HPM). Dieses war die theoretische und methodische Grundlage für drei eigene empirische Evaluationen, die in diesem Bericht dargestellt werden.

Als Vertreter der HF-Programme diente das *Critical Incident Stress Management* (CISM) für Fluglotsen. Durch die Zusammenführung von insgesamt etwa 400 Fragebogen und Interviews mit Daten des Sicherheits-, Personal- und Kapazitätsmanagements konnte nachgewiesen werden, dass der konservativ geschätzte monetäre Nutzen dieses Programms seine Kosten um den Faktor 3,5 übersteigt. Außerdem manifestierten sich weitere, intangible Nutzeneffekte z.B. eine verbesserte Organisationskultur. Diese Evaluationsstudie wird hinsichtlich Quantität und Qualität der verwerteten Daten als *Benchmark* angesehen (daher *Benchmark-Evaluation*).

Das Personalentwicklungsprogramm eines *Call Centers* der öffentlichen Verwaltung repräsentierte den HR-Bereich. Es handelte sich hierbei um ein Bündel verschiedener HR-Maßnahmen (daher Programmevaluation) wie die Einarbeitung zu Beginn der Tätigkeit, die Einweisung in Fachanwendungen, Kommunikationsschulungen, Kurzanweisungen zu aktuellen Auskunftsthemen und Teamentwicklungsmaßnahmen. Hier galt es, die Effekte der einzelnen Maßnahmen auf die Fähigkeiten/Fertigkeiten der Mitarbeiter, deren Arbeitsergebnisse und die Leistung des *Call Centers* zu isolieren. Dazu gaben alle sechs Führungskräfte und ca. die Hälfte der 43 Mitarbeiter Auskunft. Mit Hilfe eines Pfadmodells kausaler Beziehungen wurde die Auswertung strukturiert. Es konnte gezeigt werden, welche Maßnahmen in welchem Umfang mit Veränderungen von Fähigkeiten/Fertigkeiten, Arbeitsergebnissen und Organisations-*Performance* zusammenhängen und welche Optimierungsbedarfe bei den Maßnahmen bestehen. Unterschiede in der Sichtweise von Führungskräften und

Mitarbeitern gaben darüber hinaus Hinweise auf eine sinnvolle Entwicklung der Führungsarbeit.

Objekt der dritten Evaluationsstudie war das *E-learning* von Auszubildenden im Bereich der Flugsicherung. Damit ist auch der Bereich Training in den eigenen empirischen Studien vertreten. Diese Evaluation sollte mit so wenig Aufwand wie möglich (daher Kurzevaluation) eine Investitionsentscheidung bzgl. des zukünftigen Einsatzes dieser Trainingsmaßnahmen ermöglichen. Anhand einer Befragung von 32 Auszubildenden und sieben Ausbildern bzw. Ausbildungsmanagern wurde *E-learning* ökonomisch evaluiert. Der monetäre Nutzen des besten *E-learning* Moduls – bestimmt anhand des Klassen- und Simulatorunterrichts, den es ersetzen konnte – überstieg die Kosten seiner Entwicklung um das Achtfache. Andere Module haben sich dagegen noch nicht rentiert. Hier wurde ein *Break Even Point* (BEP) berechnet, der anzeigt, bei welcher Anzahl von Auszubildenden sich das Modul amortisieren wird. Außer diesen für eine Investitionsentscheidung relevanten Ergebnissen, ergaben sich viele Hinweise auf Verbesserungsmöglichkeiten der einzelnen Module bzw. ihrer Einbindung in den Ausbildungsplan.

Durch die drei Evaluationsstudien konnte das HPM als fundiertes und praktikables Rahmenmodell für Wissenschaft und Praxis bestätigt und weiterentwickelt werden. Die Erfahrungen aus den drei Evaluationsstudien sowie die Erkenntnisse der Literaturanalyse werden in einer Handlungshilfe zur ökonomischen Evaluation von Personalmaßnahmen zusammengefasst. Sie ist das Hauptprodukt des Projektes und wird Ende 2006 verfügbar sein. Als weitere Produkte sind eine Broschüre zum Thema, ein *Frequently Asked Questions* (FAQ) Instrument sowie ein Computerprogramm in Arbeit. Sie sollen die breite Öffentlichkeit für das Thema sensibilisieren und interessieren. Alle Produkte werden für den europäischen Auftraggeber EUROCONTROL – teilweise branchenspezifisch – auch in englischer Sprache abgefasst.

Schlagwörter:

Bewertung, *Call Center*, Coaching, *Controlling*, Effizienz, *E-learning*, Evaluation, Fluglotse, Humanfaktoren, Humanressourcen, Personalwesen, strategische Unternehmensführung, Stressmanagement, Training

Development and Testing of the *Human Resources Performance Model* for the Economic Evaluation of Human Factors (HF), Human Resources (HR), and Training (T) Programmes

Abstract

This document is the final scientific report of project F2105 „Guidelines for the Economic Evaluation of Human Factors HF, Human Resources HR and Training T Programmes“, which the German Federal Institute for Occupational Safety and Health launched in cooperation with the European Organisation for the Safety of Air Navigation EUROCONTROL.

Within this project, a literature review was conducted in order to identify economic evaluation methods for HF/HR/T programmes. However, neither scientific sources nor organisational practices revealed sufficiently integrated evaluation concepts that consider both, behavioural and management sciences and at the same time meet scientific and operational requirements. The economic and strategic effects of HF, HR and T in organisations have not received due attention.

The strengths of the reviewed approaches were integrated in a new framework by the name of Human Resources Performance Model (HPM). HPM was the theoretical and methodological basis of the three empirical evaluation studies reported in this document.

The Critical Incident Stress Management (CISM) for air traffic controllers represents the HF area. By combining about 400 questionnaires and interviews with data and experiences from the safety, staff, and capacity management, it could be demonstrated that the estimated fiscal benefits of the programme exceeds its costs by factor 3.5. Moreover, many intangible benefits were identified, for example an improved organisational culture. Regarding the quality and quantity of processed data, this evaluation study is considered benchmark (therefore referred to as benchmark evaluation).

The staff development programme of a public call centre represents the HR area. The programme consisted of several single interventions (therefore referred to as programme evaluation) for example ab initio training, introduction into technical applications, communication training, introduction of current contents, and team development. The challenge here was to isolate the effects of the single interventions with respect to abilities/skills, individual work output, and organisational performance. All six team leaders and about half of their 43 employees participated in the study. A path model of causal relationships structured the data analysis. It could be demonstrated, which intervention correlated to what extent with changes in abilities/skills, individual work output, and organisational performance. Moreover, improvement potential for the single interventions and the whole programme were identified. Finally, differences in the view of team leaders and team members indicated in what respects the operational leadership can be developed.

The third and final evaluation study targeted the e-learning of air traffic control ab initio trainees. Thus, also a training programme was represented in the three studies.

This evaluation was planned with the least necessary effort to facilitate an investment decision on the further use of e-learning (therefore referred to as short evaluation). Data basis was a survey with 32 trainees, seven trainers and academy managers. The fiscal benefits of the best e-learning module (estimated cost reduction by replaced of class and simulator time) exceeded the costs of the module by factor 8. However, other modules did not yet pay off monetarily. Therefore, the break even point (BEP) was calculated in form of the number of trainees necessary for the module to pay off. Apart from these results that facilitate investment decisions on the further use of e-learning, other results indicated improvement potential of single e-learning modules and their function within the curriculum.

The three evaluation studies revealed and developed the HPM as a sophisticated and feasible framework for the economic evaluation of HF/HR/T programme in science and organisational practice. Depending on the evaluation question (optimisation and extension of existing programmes, selection of new programmes, implementation and development state of the programme, resources for evaluation, stakeholder of evaluation results etc.), HPM can be used as a whole or in parts. The experiences gained in the literature review and the empirical parts of the project will be summarised in a guideline for the economic evaluation of HF/HR/T programmes. The guideline will be the main deliverable of the project and will be available by the end of 2006. Further products currently under construction are a marketing brochure, a frequently asked questions (FAQ) instrument and computer software. They should support the dissemination of the project results to a broad public. All products – partly tailored for the aviation branch – will be available in German and English.

Key words:

Air traffic controller, Assessment, Call Centre, Coaching, *Controlling*, Efficiency, *E-learning*, Evaluation, Human Factors, Human Resources, Personnel Management, Strategic Leadership, Stress management, Training

Danksagung

Die in diesem Bericht dokumentierten Ergebnisse entstanden in enger Zusammenarbeit mit den Auftraggeberorganisationen, den Praxispartnern und den dortigen Projektverantwortlichen. Unser herzlicher Dank gilt daher Dr. Birgit Köper von der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) und Dr. Manfred Barbarino von der *European Organisation for the Safety of Air Navigation* EUROCONTROL, die das Projekt initiiert (Kapitel 1.2) und fachlich stark unterstützt haben.

Die Evaluation des *Critical Incident Stress Management* (CISM, Kapitel 2.2) der Deutschen Flugsicherung (DFS) wurde erst durch die enge Zusammenarbeit und das große Engagement der Verantwortlichen in der DFS ermöglicht (Führungskräfte, Programmverantwortliche und Fluglotsen). Insbesondere möchten wir das *Monitoring* von Andreas Angenendt (Leiter Geschäftsbereich Center) sowie die Unterstützung durch Jörg Leonhardt, Thomas Makies, Frank Weber und Ralf Großböhrer hervorheben. Allen Fluglotsen, die unsere umfangreichen Fragebogen ausgefüllt haben und insbesondere denjenigen, die als kollegiale Berater CISM durchführen, gilt unser Dank.

Die Evaluation des Personalentwicklungsprogramms des teilnehmenden *Call Centers* wurde durch dessen Leiter stark unterstützt. Alle Führungskräfte des *Call Centers* und die Hälfte der Mitarbeiter füllten unseren umfangreichen Fragebogen aus, ihnen gebührt daher ebenfalls unser Dank.

Das EUROCONTROL zugehörige *Institute for Air Navigation Services* IANS in Luxemburg ist die Fachorganisation für die Evaluation der *E-learning* Programme der Flugsicherungsakademien. Der dortige Fachverantwortliche, Max Bezzina, lieferte wichtige Informationen und Impulse für diese Studie. Unsere Ansprechpartner in den teilnehmenden Akademien Brüssel und Malmö waren Johan Delauré, Marc Van Dongen-Vogels und Göran Lindqvist. Ihnen sowie den Akademieangehörigen, die unsere Fragebogen ausfüllten, sei ebenfalls gedankt.

Essen und Kopenhagen im Oktober 2006

Stefan Pennig

Joachim Vogt

1 Einleitung

1.1 Notwendigkeit ökonomischer Evaluation von Personalmaßnahmen

Die Veränderungen in Wirtschaft und Gesellschaft haben dazu geführt, dass der Bereich Personal zunehmend als zentraler Faktor im globalen und nationalen Wettbewerb gesehen wird. Entsprechend haben Unternehmen in den letzten Jahren zunehmend Anstrengungen unternommen, um ihr Personalmarketing zu professionalisieren und Programme zur Bindung wichtiger Mitarbeitergruppen² auf den Weg zu bringen.

Diese Entwicklungen legen nahe, das Human Resource (HR-) Management (HRM) in den Unternehmen organisatorisch aufzuwerten, neu zu positionieren und in den strategischen Managementprozess des Unternehmens einzubinden. Voraussetzung dafür ist allerdings, dass der HR-Bereich sich strategisch steuern lässt, wozu insbesondere eine ergebnisorientierte, ökonomische Steuerung von HR-Investitionen gehört.

Auch im praktischen Alltag eines HR-Managers, Bereichsverantwortlichen, Vorstandes oder Unternehmers gibt es vielfältige Fragestellungen, die eine gezielte Steuerung der HR-Aktivitäten erfordern. Da eine Steuerung ohne reliable und valide Messverfahren nicht möglich ist („*you can't manage what you can't measure*“³), können diese Fragen nur auf der Grundlage eines kontinuierlichen Ergebnismonitorings erfolgen, das Ergebnisse misst und bewertet (Evaluation):

- (1) Welche Personalmaßnahmen werden durch organisatorische Veränderungen erforderlich (z.B. nach einem *Merger*)? Welche Investitionen fallen dafür an? Was passiert, wenn wir hier nicht investieren?
- (2) Mit welchen personalentwicklerischen Maßnahmen sichern wir die Zukunftsfähigkeit unseres Unternehmens? Wo sollten Schwerpunkte gesetzt werden? Wie gewinnen wir die Zustimmung des Vorstandes für diese Investitionen?
- (3) Welches von mehreren alternativen Personalentwicklungsprogrammen (PE-Programmen) ist bei annähernd gleichen Kosten kurz- und langfristig nutzbringender für das Unternehmen? Falls bei den Kosten Unterschiede bestehen: Lohnt sich eine Mehrinvestition, da überproportionaler Nutzen zu erwarten ist?

² Mit Rücksicht auf die Lesbarkeit dieses Berichtes wird hier und im Folgenden für alle Personenbezeichnungen die männliche Form gewählt. Sie werden von Frauen in der weiblichen und von Männern in der männlichen Form geführt.

³ Viele der in diesem Bericht verwendeten Begriffe stammen aus dem angloamerikanischen Sprachraum. Englische Begriffe wurden daher – sofern sie auch von Deutschen häufig benutzt werden – ohne Übersetzung übernommen. Durch Kursivdruck sind sie vom deutschen Text abgehoben, um den Sprachwechsel zu erleichtern.

- (4) Welche Anforderungen an Führungskräfte ergeben sich aus der Unternehmensstrategie? Wie erheben wir den Entwicklungsbedarf in der Führungskräfteentwicklung? Wie stellen wir in den nächsten fünf Jahren fest, ob sich Investitionen in unsere Führungskräfte gelohnt haben?
- (5) Wie messen wir die Effekte eines Personalentwicklungsprogramms in der Produktion, das im Verbund mit anderen Maßnahmen (z.B. Einführung von Gruppenarbeit oder kontinuierlichen Verbesserungsprozessen KVP) durchgeführt wurde. Welche Effekte lassen sich auf die durchgeführten Trainings zurückführen?
- (6) Welchen Anteil des gesamten (Projekt-)Budgets sollten wir für PE investieren? Welcher Anteil sollte in andere Investitionen fließen, z.B. in neue Technologien? Lässt sich mit Zahlen untermauern, in welcher Maßnahme Kapital besser investiert ist?
- (7) Welche Auswirkungen und welche Bedeutung hat ggf. eine nachlassende Zufriedenheit und Arbeitsmotivation der Mitarbeiter für unseren Geschäftserfolg? Gibt es passgenaue PE-Programme, um diesen Trend aufzufangen und umzukehren? Lässt sich dies auch in einer groben Wirtschaftlichkeitsberechnung (Kapitel 2.4) darstellen?
- (8) Ist unsere Personalentwicklung strategisch aufgestellt? Wo gibt es Anpassungsbedarf? Welche positiven Effekte sind aus einer mehr strategisch orientierten Personalentwicklung für das Unternehmen zu erwarten?
- (9) Inwieweit würde unsere Wertschöpfung und Wirtschaftlichkeit in den zentralen Leistungsbereichen des Unternehmens verbessert, wenn wir ein bestimmtes PE-Programm umsetzen?
- (10) Wie gut ist das HR-Management in unserem Unternehmen und wie gut sind unsere Mitarbeiterpotenziale (Kompetenzen, Engagement) im Vergleich zu den Wettbewerbern?

Hier sind die Unternehmen und Manager auf Impulse und *Know-how* durch Fachleute angewiesen, weil sich der Personalbereich für solche ökonomischen Betrachtungen, die über ein reines Kostenmanagement hinausgehen, nicht ohne weiteres erschließt. Die Entscheidungsträger stehen hier in der Regel vor sehr komplexen Problemstellungen, die ein interdisziplinäres *Know-how* aus den Bereichen Organisationspsychologie, Arbeitswissenschaften, Personalmanagement, klassische Betriebswirtschaftslehre und Managementwissen erfordern. Die verschiedenen mit diesem Thema befassten wissenschaftlichen Disziplinen arbeiten jedoch mit sehr unterschiedlichen Ansätzen.

Gerade im HR-Bereich ist außerdem die Kluft zwischen Wissenschaft und Praxis, Personalentwicklungsexperte und Linienmanager häufig sehr groß. Die praktische Personalpolitik und Unternehmenssteuerung greift wissenschaftliche Konzepte kaum auf, weil sie auf die betriebliche Realität nicht zu passen scheinen. Dies führt dazu, dass entsprechende Steuerungskonzepte, die sowohl verhaltensbezogene als auch

ökonomische Perspektiven einbeziehen und in der Unternehmenssteuerung eingesetzt werden können, kaum genutzt werden.

Diese wissenschaftliche Arbeit versteht sich als interdisziplinärer Beitrag zum HR-Management, der verhaltens- und betriebswirtschaftliche Perspektiven näher zusammennückt. Damit wird eine funktionale Brücke über die oben beschriebene Kluft zwischen Unternehmensführung und HRM in Wissenschaft und Praxis geschlagen.

1.2 Entstehung des Projekts

Die in Kapitel 1.1 dargestellten Überlegungen wurden für die Luftfahrtbranche auf dem *Workshop „Interfaces in Air Traffic Management“* im Dezember 2002 in Dortmund thematisiert. Vertreter der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin BAuA bzw. der *European Organisation for the Safety of Air Navigation* EUROCONTROL initiierten daraufhin 2003 eine Machbarkeitsstudie (Kapitel 2.2.2). Diese sollte klären, ob eine Evaluation von Maßnahmen im Personalbereich mit ökonomischen Methoden überhaupt möglich und sinnvoll ist. Als Zielprogramm diente das *Critical Incident Stress Management* (CISM) der Deutschen Flugsicherung (DFS). In enger Zusammenarbeit mit dem verantwortlichen Programmmanager der DFS wurde eine Kombination zweier Evaluationsansätze aus den beteiligten Disziplinen angewendet: Die Humanressourcenevaluation der Organisationspsychologie und die Nutzwertmethode der Wirtschaftswissenschaften. Hauptergebnis der Machbarkeitsstudie war, dass diese Methoden sinnvoll auf das CISM-Programm bezogen eingesetzt werden konnten. Neben einer Vielfalt intangibler Nutzeneffekte konnte eine konservative Schätzung des *Return on Investment* (ROI) des CISM-Programms von 250% nachgewiesen werden (Vogt, Leonhardt, Köper & Pennig, 2004).

Aufgrund der viel versprechenden Ergebnisse der Machbarkeitsstudie (Kapitel 2.2.2) schrieb EUROCONTROL im Jahr 2004 zusammen mit der BAuA eine Hauptstudie zu diesem Thema aus. Nach einer Sichtung vorhandener Praxiskonzepte und wissenschaftlicher Untersuchungen zum Thema (Pennig et al., 2006), sollten mindestens drei weitere Programme aus den Bereichen Humanfaktoren (HF), Humanressourcen (HR) und Training (T) evaluiert werden. Aufgrund der Literaturlauswertung und der eigenen empirischen Erfahrung mit den unterschiedlichen Bereichen und methodischen Aufwendungen werden Handlungsempfehlungen für die ökonomische Evaluation von Personalmaßnahmen abgeleitet. Diese sind das Hauptergebnis des Projektes und werden Ende 2006 zugänglich sein. Die einzelnen Produkte des Projekts und ihre Funktion in der Gesamtkonzeption werden im Folgenden kurz dargestellt.

1.3 Produkte des Projekts

Die zum Thema bereits vorliegende Literatur wurde in einer Literaturübersicht zusammengefasst (Pennig et al., 2006). Diese ist unter dem Menüpunkt „Informationen für die Praxis“ auf der *Homepage* der BAuA sowie als Buch der Schriftenreihe der BAuA verfügbar und wird im folgenden Kapitel 1.3.1 kurz dargestellt.

Ein wichtiges Ergebnis der Literaturrecherche war, dass zwar vielfältige Methoden der ökonomischen Evaluation von HF/HR/T-Maßnahmen zur Verfügung stehen, diese aber wenig integriert sind. Auf wissenschaftlicher Seite war insbesondere die Integration von Ansätzen der Wirtschaftswissenschaften (z.B. *Balanced Scorecard*, Kaplan & Norton, 1996) mit solchen der Organisationspsychologie (z.B. Trainingseva-

uation, Kirkpatrick, 1994) bisher wenig fortgeschritten. In der Praxis besteht erhebliches Potential in der Integration der im Unternehmen ohnehin vorhandenen Daten z.B. in den Bereichen *Controlling* oder Qualitätsmanagement zum Zwecke der Steuerung und Evaluation von HF/HR/T-Maßnahmen. Auf der Grundlage der Literaturanalyse und der im vorliegenden Bericht dokumentierten eigenen empirischen Untersuchungen (Kapitel 1.3.2 und 2) wurde daher ein Rahmenkonzept zur Planung, Steuerung und Evaluation von HF/HR/T-Maßnahmen entwickelt, das die Stärken der verschiedenen Ansätze aus Wissenschaft und Praxis vereint (*Human Resources Performance* Modell, HPM, Kapitel 1.4).

Die weiteren Produkte des Projekts bauen auf diesem Rahmenkonzept auf. Nachdem die Literaturübersicht und die eigenen empirischen Untersuchungen eher wissenschaftlichen Zwecken dienen, werden die weiteren Produkte stärker auf die Bedürfnisse der Praxis zugeschnitten sein. Die Handlungshilfe (Kapitel 1.3.3) wird dem Praktiker eine schnelle Orientierung und praktische Umsetzung einer Evaluationsstudie ermöglichen. Sie wird ergänzt werden um ein Computerprogramm (*Economic Evaluation Selection Tool* EMS), das die Auswahl und den Einsatz von Evaluationsmethoden unterstützen soll (Kapitel 1.3.4). *Frequently Asked Questions* (FAQ) und *Marketing* im Bereich der ökonomischen Evaluation von Personalmaßnahmen werden aufbereitet mit dem Ziel, dieses Thema zu verbreiten (Kapitel 1.3.5). Alle Produkte werden für den Auftraggeber EUROCONTROL, teilweise branchenspezifisch für die Luftfahrt, auch in englischer Sprache abgefasst.

1.3.1 Literaturübersicht

Methoden der Steuerung und ökonomischen Evaluation von HF/HR/T-Maßnahmen wurden in der Literaturübersicht hinsichtlich ihrer Theorie und Methodik zusammenfassend dargestellt. Anwendungen zu diesen Verfahren wurden recherchiert und als Fallbeispiele beschrieben, sofern sie öffentlich zugänglich waren. Praktische Verwendbarkeit, Schwächen und Stärken der Verfahren sind in der Literaturübersicht dargestellt. Ein die Stärken der verschiedenen Ansätze vereinendes Rahmenkonzept zur Planung, Steuerung und Evaluation von HF/HR/T-Maßnahmen, das *Human Resources Performance Modell* (HPM, Kapitel 1.4) wurde neu entwickelt. Es orientiert sich am Zweck und dem Adressaten der Evaluationsergebnisse, die im Folgenden kurz aufgeführt sind:

Kapitalgeber: Die Bewertung des Humankapitals im Sinne eines Werteausweises ist mit den gängigen Modellen der Bilanzierung wenig aussagekräftig. Mit Hilfe ausgewählter, nicht unbedingt ökonomischer Indikatoren lässt sich besser erfassen und verstehen, welche aktuellen und zukünftigen Wettbewerbsvorteile durch die bestehenden Management- und Mitarbeiterkompetenzen erwachsen.

Top-Management: *Scorecards*, die das Personal als wichtigen Wertetreiber im Unternehmen abbilden, helfen, die grundlegenden Bewertungskriterien guter Personalarbeit im Sinne des Unternehmens zu definieren und die wichtigsten, strategieumsetzenden Handlungsfelder zu erkennen. Entscheidungshilfen und Erfolgskriterien für einzelne HF/HR/T-Maßnahmen müssen maßnahmenbezogen mit dem Personalbereich erarbeitet werden.

Bereichsmanagement: *Bereichsmanager* können z.Zt. auf allgemeine, strategische Steuerungskonzepte zurückgreifen, um diese für den eigenen Verantwortungsbereich

zu spezifizieren. Die Ausrichtung der Personalentwicklung auf die Unterstützung der Leistungsprozesse sollte im engen Dialog von Fach- und Führungskräften, HR-Verantwortlichen, Prozessverantwortlichen und Mitarbeitern erfolgen, um klare Ziele und Bedarfe zu erarbeiten. Spezifische Modelle in diesem Bereich fehlen.

Operative Führungsebene: Teamleiter können vor allem klassische Konzepte der Aufgaben- und Anforderungsanalyse nutzen, um den Entwicklungsbedarf zu erheben. Dieser ist zugleich die Messlatte für die Personalentwicklungsarbeit.

Personalabteilung: HF/HR/T-Manager haben vielfältige Möglichkeiten, um die eigene Arbeit zu steuern und zu evaluieren. Einerseits ist die Mitwirkung an einer unternehmensweiten *Scorecard* sinnvoll, um den Einfluss und die Bedeutung der Personalentwicklung strategisch zu verankern. Andererseits kann eine eigene Personal-*Scorecard* daraus entwickelt werden, die das Leistungsspektrum des Personalbereiches differenzierter erfasst. Schließlich können die Methoden der traditionellen HR-Evaluation genutzt werden, um die Planung und die Kontrolle einzelner HR-Programme zu systematisieren.

Die Literaturanalyse hat gezeigt, dass wissenschaftliche und praktische Ansätze zur Steuerung und Evaluation von HF/HR/T-Maßnahmen bisher wenig integriert sind. Es wurden daher eigene empirische Untersuchungen durchgeführt, um einen integrativen Ansatz zu entwickeln und dessen Nützlichkeit für Forschung und Praxis zu erproben.

1.3.2 Eigene empirische Untersuchungen

Mit dem oben erwähnten und in Kapitel 1.4 näher beschriebenen *Human Resources Performance* Modell als theoretischer und methodischer Grundlage wurden drei empirische Evaluationsstudien durchgeführt. Die Zahl drei ergab sich, weil *Personalmanagement* in drei große Bereiche gegliedert werden kann, die jeweils mit einer eigenen Studie repräsentiert sein sollten. *Human Resources Management* (HRM) ist die branchenübergreifende Bezeichnung für alle Aktivitäten, die menschliche Arbeitskraft zur Verfügung stellen, weiterentwickeln und erhalten, also z.B. Selektion und Gesunderhaltung von Personal. *Human Factors* (HF) ist ein etablierter Bereich insbesondere in sicherheitskritischen Branchen wie Transport- und Energiewirtschaft. HF entwickelte sich in der Folge von Großschadensereignissen wie Verkehrs- und Kraftwerksunfällen, die mit so genanntem „menschlichem Versagen“ erklärt wurden. Die treffendere Bezeichnung hierfür ist jedoch der von James Reason (1986) eingeführte Begriff des „*organisational accident*“, der zum Ausdruck bringt, dass nicht der Mensch versagt, sondern die Organisation, die es versäumt hat, den *Human Factor* hinreichend zu berücksichtigen. *Training* (T) schließlich ist der Teilbereich von HRM, der sich speziell mit der Ausbildung von Personal beschäftigt. Er nimmt insbesondere in den Wirtschaftszweigen viel Raum ein, die auf hoch qualifizierte Spezialisten angewiesen sind, wie beispielsweise Piloten oder Fluglotsen.

Als HF-Programm diente das bereits erwähnte CISM-Programm der DFS, für das eine wissenschaftlich belastbare *Benchmark-Evaluation* durchgeführt werden sollte (Kapitel 2.2). Vertreter des Bereiches HR wurde ein Bündel von Personalmaßnahmen (Programmevaluation, Kapitel 2.3) für *Call Center* Agenten; dieses sollte mit mittlerem Aufwand evaluiert werden, der auch von einem Unternehmen leistbar ist, das nicht über Forschungsmittel verfügt. Als drittes Zielprogramm wurde *E-learning*

für Fluglotsenanwärter als Trainingsmaßnahme ausgewählt (Kapitel 2.4). Der Evaluationsansatz folgte hier dem Paretoprinzip, d.h. mit minimalem Aufwand sollten Daten generiert werden, die dem Management der Flugsicherungsakademien eine Investitionsentscheidung im Bereich *E-learning* ermöglichen (Kurzevaluation).

1.3.3 Handlungshilfe

Die Handlungshilfe wird geschrieben für Fachverantwortliche, Experten und Berater aus den Bereichen Personal, betriebliche Gesundheitsförderung (BGF), Arbeitssicherheit und Arbeitsschutz sowie an Manager und Führungskräfte in Organisationen, die Personalverantwortung tragen und/oder die Personalführung und –entwicklung in ihrem Bereich optimieren wollen⁴. Sie soll Hilfestellungen für das Management von Personalmaßnahmen in der Praxis geben, insbesondere wenn es darum geht, Wirtschaftlichkeitsbetrachtungen in die Planung, Steuerung, Bewertung und/oder Kontrolle von HF/HR/T-Maßnahmen einzubeziehen. Die Handlungshilfe möchte dazu beitragen, dass die traditionelle, verhaltenswissenschaftliche Evaluation im Bereich des Personalmanagements um die betriebswirtschaftliche Perspektive erweitert wird und Evaluation als Gemeinschaftsprozess von Verhaltensexperten, Ökonomen und Bereichsverantwortlichen in Betrieben verstanden wird. Zu diesem Zweck wird die Handlungshilfe bieten:

- eine Landkarte zur Strukturierung und Systematisierung der ökonomischen Evaluation von Personalmaßnahmen,
- Vorgehensmodelle für die Planung, Steuerung, Bewertung und Kontrolle von Personalmaßnahmen, sowie
- einen Werkzeugkasten (*Toolbox*) mit professionellen Instrumenten (*Tools*) zur ökonomischen Evaluation.

Die Handlungshilfe wird sowohl eine umfassende Orientierung in diesem komplexen Handlungsfeld als auch sehr konkrete Anleitungen für ausgewählte Fragestellungen ermöglichen, z.B. wie ein Fragebogen und ein Untersuchungsdesign erstellt werden können. Die Handlungshilfe wird sich in vier Themenblöcke gliedern (Tabelle 1). Zur schnellen Orientierung werden am Ende eines jeden Kapitels *Tools* und weiterführende Literatur eingefügt.

⁴ Wenn in diesem Bericht von Personalmanagement und Personalwesen gesprochen wird, sind jeweils alle angesprochenen Fachbereiche (Personalbetreuung, Personalentwicklung, BGF, Arbeitsschutz, Arbeitssicherheit) gemeint, die in der Praxis häufig auch dem Personalwesen zugeordnet sind.

Tabelle 1: Gliederung der Handlungshilfe

| Kapitel | Inhalt |
|---|---|
| Einführung | Kapitel 1 führt in das Thema ein und orientiert über Nutzung und Struktur der Handlungshilfe gegeben. |
| Grundlegende Handlungshilfen zur Planung und Strukturierung von Evaluationsvorhaben | Das Kapitel 2 zeichnet eine Landkarte zur ökonomischen Evaluation. Sie dient als Planungshilfe und zur Strukturierung des Vorgehens in der Evaluation unabhängig von der konkreten Fragestellung. |
| Spezielle Handlungshilfen für ausgewählte Fragestellungen getrennt nach Anwendern bzw. Nutzergruppen der Evaluation | Die Kapitel 3-6 bieten Informationen und Werkzeuge für ausgewählte Zielsetzungen in der ökonomischen Evaluation. In den jeweiligen Kapiteln werden vier verschiedene Anwender- bzw. Nutzergruppen von Evaluation angesprochen. Der Leser hat so die Möglichkeit, gezielt ausgewählte Abschnitte der Handlungshilfe zu nutzen. |
| Ergänzende Hinweise und Informationen | Im Kapitel 7 finden sich ergänzende Informationen zur wissenschaftlichen Arbeitsmethodik in der Evaluation, z.B. zur Gestaltung von Erhebungsinstrumenten. |

1.3.4 *Economic Evaluation Selection Tool EMS*

Die als Buch, Bericht oder Broschüre herausgegebenen Erkenntnisse des Projekts sollen ergänzt werden um das Computerprogramm EMS, das die Auswahl und Umsetzung von Evaluationsmethoden in der Praxis unterstützt. Derzeitiger Planungsstand ist ein wordbasiertes Programm auf der Grundlage der in diesem Bericht dargestellten Struktur- und Prozessmodelle (Kapitel 1.4). Richtungweisend sind z.B. die Zielgruppen (Kapitel 1.3.1), denen der Anwender angehört, also beispielsweise Bereichsmanager oder operative Führungskraft. Weiter wird die Vorgehensweise bestimmt durch die Art der Fragestellung, etwa Evaluation einer vorhandenen HF/HR/T-Maßnahme oder Auswahl zukünftiger HF/HR/T-Maßnahmen. Auch der geplante Aufwand der Evaluation sowie die bereits verfügbaren Daten (z.B. aus *Controlling*, Qualitäts-, Kapazitäts- und *Personalmanagement*) werden vom EMS berücksichtigt. Ein Prototyp des Computerprogramms wird Ende 2006 verfügbar sein.

1.3.5 *Frequently Asked Questions FAQ und Marketing*

In dem vorliegenden Projekt sollten nicht nur Methoden der Steuerung und Evaluation von HF/HR/T-Maßnahmen entwickelt, sondern auch Instrumente zur professionellen Verbreitung des Themas entwickelt werden. Zwei Instrumente zur Sensibilisierung der Öffentlichkeit für dieses Thema werden erstellt: Ein *Frequently Asked Questions* FAQ Instrument für das Informationszentrum der BAuA und eine Informationsbroschüre. Das FAQ-Instrument wird erarbeitet für Personengruppen (z.B. HR-Manager), die die Relevanz der Steuerung und Evaluation von HF/HR/T-Maßnahmen für ihre Arbeit bereits entdeckt haben und daher in der BAuA anrufen, um weitere Informationen zu erhalten. Ein Wissensdatensatz mit einer kurzen Problembeschreibung, Schlagworten und der bereits verfügbaren Literaturübersicht (Kapitel 1.3.1) sowie den in Arbeit befindlichen Dokumenten (Kapitel 1.3.2-1.3.4) liegt dem Informationszentrum der BAuA seit Juni 2006 vor. Auch die *Marketing*-Broschüre richtet sich

an diese Personen; sie ist aber darüber hinaus für Zielgruppen gedacht, die das Thema noch nicht entdeckt haben.

1.4 Das *Human Resources Performance* Modell HPM

Das *Human Resources Performance* Modell (HPM) ist die theoretische und methodische Grundlage der in diesem Bericht dokumentierten empirischen Evaluationsstudien. Das HPM kann verstanden werden als Rahmenkonzept zur Steuerung und Evaluation von Personal- und Organisationsentwicklungsmaßnahmen (Auswahl von Mitarbeitern, Qualifizierungsprogramme, Maßnahmen zur Förderung von Gesundheit und Arbeitsmotivation, Programme zur Führungskräfteentwicklung etc.).

Das HPM geht von folgenden Überlegungen aus: Eine Evaluation versorgt Entscheidungsträger mit Informationen zur Wirksamkeit von Maßnahmen. Die Informationen dienen vor allem der Entscheidungshilfe bzw. der Steuerung von Maßnahmen. Solche Steuerungsinformationen werden bei der Planung, Umsetzung und Kontrolle von Maßnahmen benötigt. Daher sollte eine Evaluation als ein *Controlling* verstanden werden, das die Gestaltung, Implementierung und Optimierung von Maßnahmen unterstützt. In diesem Sinne unterstützt eine Evaluation oder das *Controlling* die Steuerung von Personalmaßnahmen mit Informationen von ihrer Planung über ihre Umsetzung und ggf. bis hin zu ihrer Absetzung.

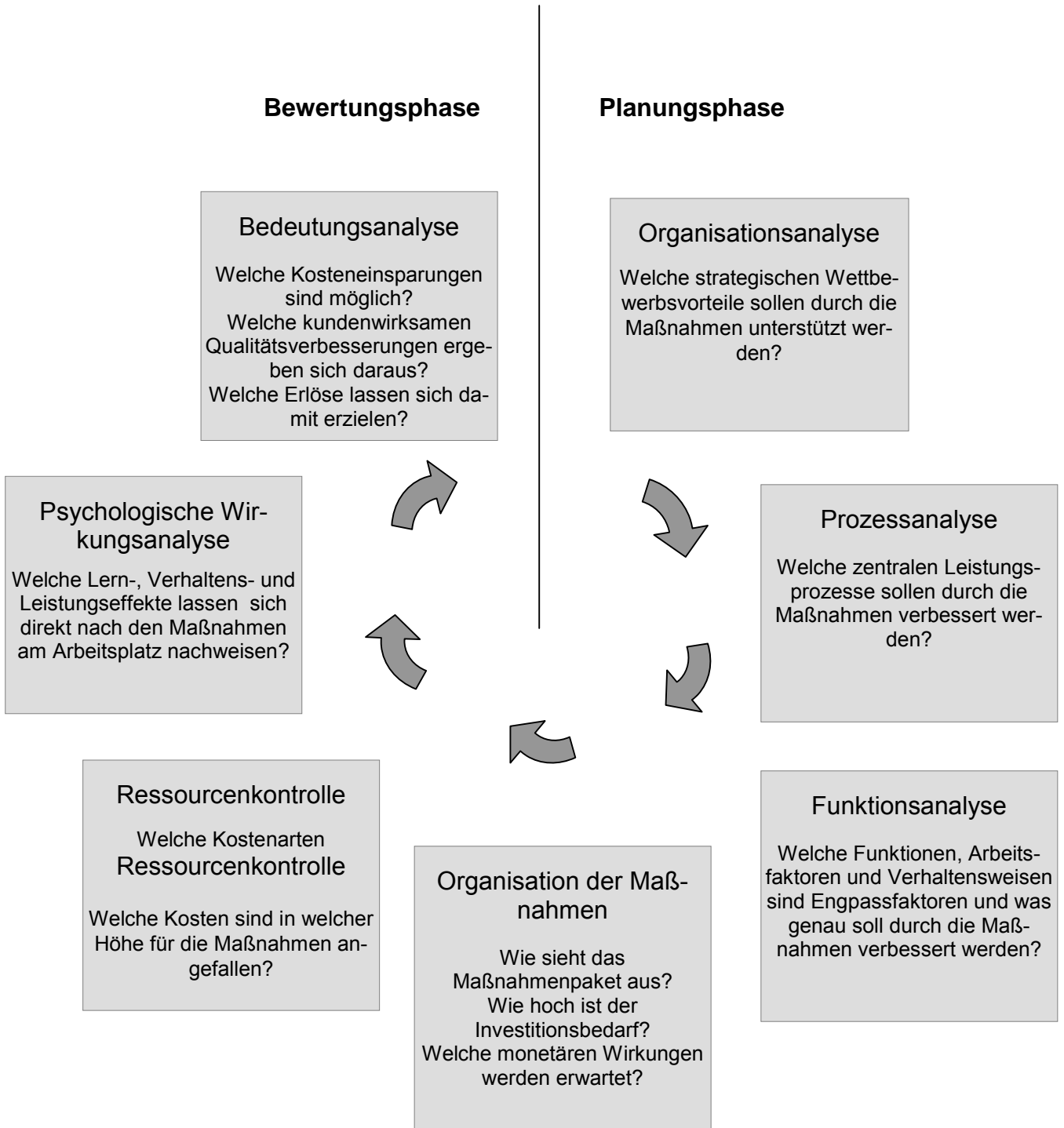
Idealerweise beginnt der Evaluationsprozess also im Rahmen der Planung von Maßnahmen. Wichtige Teilprozesse einer Evaluation, wie Zielklärung, Operationalisierung der Erfolgskriterien und Ableitung von Messverfahren sind Bestandteil der Planung bzw. unterstützen diese. Außerdem ist es in manchen Fällen sinnvoll, vor der Programmplanung und Implementierung eine Basismessung durchzuführen (z.B. eine Bildungsbedarfsanalyse bevor ein *Training* aufgelegt wird), um eine Feinjustierung der Ziele und Maßnahmen und eine Vergleichskontrolle vornehmen zu können. Darüber hinaus dient Evaluation in vielen Fällen auch der Optimierung von Maßnahmen im laufenden Umsetzungsprozess.

Die Evaluation von HF/HR/T-Maßnahmen in Organisationen erfordert die Betrachtung unterschiedlicher Wirkungsebenen, die im HPM berücksichtigt sind:

- (1) die Individual- oder Funktionsebene: Auswirkungen auf den einzelnen Mitarbeiter
- (2) die Prozessebene: Auswirkungen auf das „gemeinsame Arbeiten“
- (3) die Organisationsebene: Auswirkungen auf die „Überlebensfähigkeit“ der Organisation

Organisationen entscheiden sich für HF/HR/T-Maßnahmen, wenn sie auf allen drei Ebenen wirksam sind. Viele traditionelle Evaluationen beschränken sich jedoch nur auf die Individualebene. Für eine Organisation ist es aber bedeutsam, ob die durch die Maßnahme verbesserte Qualifikation, Arbeitsmotivation oder Fitness sich in höherer Qualität oder geringeren Kosten des Leistungsprozesses niederschlagen. Dies wird jedoch erst deutlich, wenn man die Prozess- und Organisationsebene in den Evaluationsprozess einbezieht.

Das HPM unterscheidet insgesamt sieben Arbeitsschritte der Evaluation beginnend mit der Organisationsanalyse und endend mit der ökonomischen Bedeutungsanalyse (s. Abbildung 1). Die sieben Schritte werden im Folgenden erläutert.

Abbildung 1: Das *Human Resources Performance* Modell

Schritt 1: Organisationsanalyse

Letztlich bestimmt sich der Nutzen einer HF/HR/T-Maßnahme aus dem Beitrag, den sie zur Umsetzung der Organisationsstrategie erbringt. Die Wirtschaftlichkeitsanalyse beginnt daher mit der Klärung der Organisationsstrategie. Eine Organisationsstrategie ergibt sich aus der Einschätzung des aktuellen und zukünftigen Marktes, in dem die Organisation auftritt, aus den Wettbewerbsfaktoren und Kundenerwartungen und den vorhandenen und erreichbaren Wettbewerbsvorteilen des Unternehmens in diesem Markt.

Die Organisationsstrategie umfasst grundlegende strategische Zielsetzungen (z.B. Rentabilität, Umsatzwachstum, Organisations- und Produktimage, angestrebte Wettbewerbsvorteile etc.) und strategische Handlungsfelder (Verbesserung im Bereich *Service*, Entwicklung innovativer Produkte, Aufbau eines neuen Vertriebskanals etc.). Der Wirkungsgrad einer HF/HR/T-Maßnahme bestimmt sich über den positiven Einfluss auf die strategischen Ziele und Handlungsfelder. Um diesen Einfluss abzubilden und damit die Wirtschaftlichkeit einer HF/HR/T-Maßnahme auf der Organisationsebene nachzuweisen, müssen Indikatoren für die Organisationsstrategie gebildet werden.

Ohne definierte Organisationsstrategie lässt sich die Bedeutung einer Maßnahme für die Organisation nicht einschätzen. Wird eine HF/HR/T-Maßnahme ohne Bezug zur Organisationsstrategie geplant und evaluiert, bleibt die Bedeutung verhaltensbezogener und monetärer Wirkungen unklar. Ob Kosteneinsparungen, Produktverbesserungen oder Kompetenzentwicklungen: jegliche Veränderung ist erst aus dem Blickwinkel der Strategie für eine Organisation wertvoll, neutral oder schädlich. Daher bedeutet dieser Planungsschritt für das HF/HR/T-Management:

- HF/HR/T-Maßnahmen werden von Anfang an aus der Organisationsstrategie abgeleitet.
- Die Evaluation wird an strategischen Zielen und entsprechenden Indikatoren ausgerichtet.
- Der Evaluationsprozess beginnt bereits hier mit der Festlegung von strategischen Messkriterien und der Entwicklung entsprechender Messverfahren.

Wichtige Fragestellungen im Rahmen der Organisationsanalyse sind:

- Welche Organisations-*Performance* wird zukünftig entscheidend sein, um die Überlebensfähigkeit der Organisation zu sichern?
- Wo sind zukünftig attraktive Märkte und wo und in welcher Form kann Überlegenheit erreicht werden?
- Welche besonderen Ressourcen sind dazu in der Organisation verfügbar und welche müssen ausgebaut werden?
- Welche strategischen Ziele und Handlungsfelder ergeben sich daraus?
- Durch welche Indikatoren lassen sich Ziele und Handlungsfelder abbilden?
- Welche strategischen Ziele und Handlungsfelder werden durch die geplante HF/HR/T-Maßnahme beeinflusst?

Schritt 2: Prozessanalyse

Im zweiten Schritt werden die Leistungsprozesse in der Organisation betrachtet. Leistungsprozesse sind die Einheiten, in denen die Wertschöpfung stattfindet und Kosten in Abhängigkeit von der Leistungsmenge entstehen. Der Nutzen einer HF/HR/T-Maßnahme resultiert aus dem positiven Einfluss auf diese Leistungsprozesse. Solche positiven Einflüsse können in Effizienzsteigerungen, Kosteneinsparungen oder einer verbesserten Qualität liegen.

Um die Wirksamkeit einer HF/HR/T-Maßnahme auf der Prozessebene nachzuweisen, werden Indikatoren für Qualität und Effizienz der relevanten Leistungsprozesse benötigt: Beispiele sind Produktqualität, Prozesskosten, Durchlaufzeiten, Qualität der internen Kunden-Lieferanten-Beziehungen, Qualität der Informationsprozesse.

In der Planung der HF/HR/T-Maßnahme bedeutet dies:

- Die Maßnahme wird auf die strategisch relevanten Leistungsprozesse und die Optimierung von Prozessfaktoren ausgerichtet.
- Für die Evaluation werden Prozessziele definiert, die die Wirkung der Maßnahme abbilden.
- Der Evaluationsprozess wird hier fortgesetzt mit der Operationalisierung von prozessualen Messkriterien und Messverfahren.

Wichtige Fragestellungen im Rahmen der Prozessanalyse sind:

- Welche Leistungsprozesse müssen im Fokus der Entwicklungsmaßnahmen stehen?
- Wo werden Wertschöpfung, Kundenzufriedenheit, Umsätze oder Deckungsbeiträge mittel- und langfristig bewirkt oder erwirtschaftet?
- Wo sind Effizienzverbesserungen notwendig?
- Welche Indikatoren bilden die geplante Prozessqualität und –effizienz ab?
- Welche Prozessmerkmale werden durch die geplante HF/HR/T-Maßnahme beeinflusst?

Schritt 3: Funktionsanalyse (Aufgaben- und Verhaltensanalyse)

Der Aufgaben- und Verhaltensbereich ist das Feld traditioneller HF/HR/T-Evaluation. Hier geht es um die Wirkung von Maßnahmen auf das Verhalten, die Leistungsfähigkeit und die Leistung des einzelnen Mitarbeiters.

Im Planungsprozess geht es darum, präzise zu definieren, welche die Leistungsfähigkeit bestimmende Faktoren im Fokus der Entwicklungsmaßnahmen stehen sollten: die Arbeitsumgebung, die Arbeitsmittel, die individuelle Qualifikation, die Einstellungen und Motivation der Mitarbeiter, die Gesundheit und Fitness, die Führung, die Zusammenarbeit und das Teamverhalten, die Informationsprozesse, die Arbeitsabläufe, die Regelung von Verantwortlichkeiten, die Vergütungsstruktur etc.

Die Planungen für eine HF/HR/T-Maßnahme und deren Evaluation werden also in diesem Schritt auf die Förderung und Überprüfung des individuellen Leistungsprozesses ausgerichtet. Dafür müssen auch hier Indikatoren und Messgrößen sowie reliable und valide Messverfahren entwickelt werden. Dies bedeutet:

- Die HF/HR/T-Maßnahme wird auf die strategisch relevanten individuellen Leistungs- und Verhaltensziele ausgerichtet.
- Für die Evaluation werden Leistungs- und Verhaltensziele definiert, die die Wirkung der Maßnahme abbilden.
- Der Evaluationsprozess wird fortgesetzt mit der Operationalisierung von leistungs- und verhaltensbezogenen Messkriterien und Messverfahren.

Schritt 4: Organisation der Maßnahmen (Interventionsplanung und Investitionsrechnung)

Die systematische Analyse auf Organisations-, Prozess- und Funktionsebene gewährleistet, dass die geplante Maßnahme möglichst passgenau an den erfolgskritischen und die Leistung bestimmenden Faktoren ansetzt. In der Regel sind es nicht einzelne Ursachenfaktoren, die für die *Performance* entscheidend sind und isoliert gefördert werden sollten. In dieser Phase gilt es, die Interventionen im Verbund zu planen, die daraus resultierenden Investitionskosten zu berechnen und eine Prognose der monetären Wirkungen im Sinne einer Investitionsrechnung zu erstellen.

Die zentralen Fragestellungen der Interventionsplanung und Investitionsrechnung sind:

- Wie sieht ein Maßnahmenpaket aus, das an den verschiedenen zu entwickelnden Faktoren ansetzt und vor allem die Erfolgsfaktoren auf Prozess- und Funktionsebene im Blick hat?
- Welche Kosten entstehen und wie hoch ist der Investitionsbedarf für die Maßnahme pro Jahr?
- Welche Kosteneinsparungen und Erlöse ergeben sich aus den Wirkungen auf Prozess- und Organisationsebene?
- Wie und wann werden die verschiedenen Messverfahren begleitend zur Umsetzung der HF/HR/T-Maßnahmen eingesetzt?

Schritt 5: Ressourcenkontrolle (Kostenrechnung)

Beginnend mit der Umsetzung der HF/HR/T-Maßnahme werden die anfallenden Kosten erfasst. Eine begleitende Kostenrechnung berücksichtigt alle direkt und indirekt zurechenbaren Kostenarten: Personalkosten, Reisekosten, Sachkosten, Honorarkosten etc. Es gilt, alle Kosten systematisch zu erfassen und im Rahmen einer Kostenträgerrechnung den einzelnen Elementen des HF/HR/T-Maßnahmenpaketes (z.B. Kosten pro Trainingstag) oder dem einzelnen Teilnehmer des Programms (Kosten pro Teilnehmer) zuzuordnen.

Die zentralen Fragestellungen in diesem Feld sind:

- Welche Kosten pro Kostenart sind angefallen?
- Welche fixen und welche variablen Kosten sind bisher entstanden?
- Was kostet eine weitere Maßnahme (z.B. die Durchführung eines weiteren Trainings)?

Schritt 6: Psychologische Wirkungsanalyse

Die psychologische Wirkungsanalyse umfasst die gesamte Datenerhebung und Auswertung auf Verhaltens- und Aufgabenebene (Funktionsebene). Untersucht werden folgende Fragen:

- Welche Veränderungen in der Leistungsfähigkeit, im Verhalten und in der Leistung haben sich aus der Maßnahme ergeben?
- Auf welche Ursachenfaktoren sind diese zurückzuführen (z.B. erworbene Fertigkeiten, veränderte Einstellungen, bessere Informationsprozesse)?

Untersucht werden damit drei individuelle Wirkungsebenen:

- (1) der Einfluss auf die Kompetenzen (Wissen, Fertigkeiten, Gesundheit und Fitness, Fähigkeiten, Einstellungen und Werte)
- (2) der Einfluss auf das Verhalten am Arbeitsplatz (Qualität mit der die relevanten Tätigkeiten am Arbeitsplatz ausgeübt werden)
- (3) der Einfluss auf die individuelle Leistung (Quantität und Qualität der erzielten Ergebnisse)

Wichtige methodische Fragestellungen der psychologischen Wirkungsanalyse sind:

- Wie lassen sich die Wirkungen zuverlässig und genau erfassen?
- Wie lassen sich die Wirkungen nachweisbar der Maßnahme zuordnen?
- Wie lassen sich Zusammenhänge zwischen Kompetenzen, Verhalten und Leistung aufzeigen?

Schritt 7: Ökonomische Bedeutungsanalyse

Die abschließende ökonomische Bedeutungsanalyse soll aufzeigen, welche Wirkungen diese Veränderungen von Kompetenzen, Verhalten und individueller Leistung auf der Prozess- und Organisationsebene haben.

Ökonomisch bedeutsam sind Verhaltens- und Leistungsveränderungen, die die Kosten und/oder die (erlöswirksame) Qualität der Geschäftsprozesse positiv beeinflussen. Die Steuerung von Kosten und Qualität der Geschäftsprozesse findet im Prozessmanagement statt. Die folgende Grafik veranschaulicht den Einfluss von individuellen Verhaltens- und Leistungsveränderungen auf das Prozessmanagement und im Zusammenspiel mit anderen Einflussbereichen (s. Abbildung 2). Folgende Zusammenhänge sind von Bedeutung:

Individuelles Verhalten kann die Qualität eines Produktes oder einer Dienstleistung direkt (z.B. im direkten Kundenkontakt) oder indirekt (z.B. im Zusammenspiel mit anderen Personen) beeinflussen. Ökonomisch bedeutsam sind Qualitätsverbesserungen, die z.B. die Kundenzufriedenheit, die Marktposition, den Umsatz (Absatzmenge, erzielbarer Marktpreis) und damit den Deckungsbeitrag fördern.

Individuelles Verhalten kann die kalkulatorischen Kosten zur Herstellung eines Produktes oder einer Dienstleistung beeinflussen, indem z.B. durch weniger „Ausschuss“ in der Fertigung weniger Material verbraucht wird (geringere Materialkosten) oder pro Stunde mehr Produkte oder Dienstleistungen produziert werden (geringere Personalkosten, geringere Maschinenkosten).

Ob diese Qualitäts- und Kosteneinsparpotenziale zu Erlösen, realen Kosteneinsparungen und Renditen führen, hängt von einer Reihe weiterer Steuerungsmaßnahmen ab.

Das Qualitätsmanagement trägt dafür Sorge, dass Verhaltensverbesserungen im Sinne der Kundenerwartungen genutzt werden.

Das Kapazitätsmanagement bewirkt, dass durch Verhaltens- und Leistungsverbesserungen frei werdende Kapazitäten in der Organisation erkannt und erlöswirksam woanders genutzt oder kostenwirksam eingespart werden. Für die gezielte Steuerung der Leistungskapazität liefert die Prozesskostenrechnung die entscheidungsrelevanten Informationen.

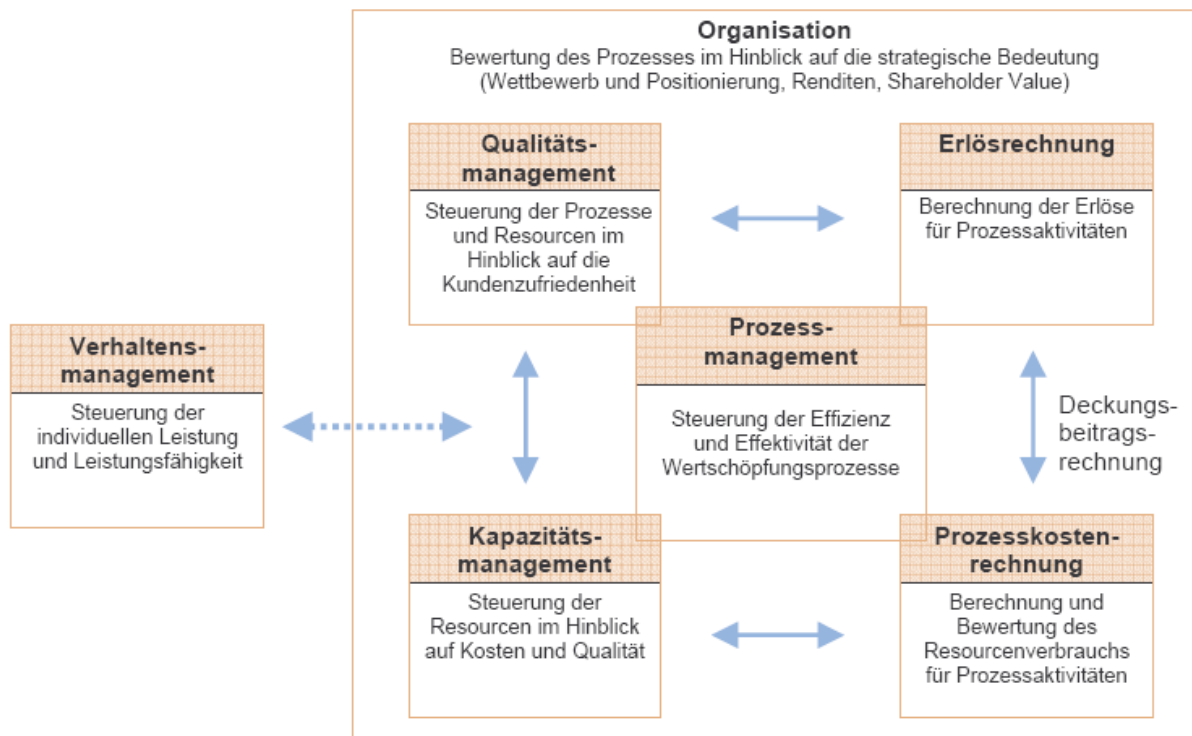


Abbildung 2: Einbindung des Verhaltensmanagements in das Prozessmanagement

Eine ökonomische Bewertung von HF/HR/T-Maßnahmen erfordert, das Zusammenspiel mit anderen *Managementsystemen* im Unternehmen, insbesondere dem Qualitäts- und dem Kapazitätsmanagement, zu betrachten. Damit können kalkulatorische und reale Kosten- und Erlöseffekte differenziert betrachtet werden.

Die zentralen Fragestellungen in diesem Feld sind:

- Welche Qualitäts- und Kapazitätseffekte ergeben sich aus den Wirkungen der HF/HR/T-Maßnahme?
- Wie werden diese genutzt oder wie könnten sie genutzt werden?
- Welche monetären Effekte im Bereich der Prozesskosten und der Erlöse am Markt ergeben sich daraus?
- Welche Rendite resultiert aus diesen Effekten für die HF/HR/T-Maßnahmen im betrachteten Zeitraum?

1.5 Zusammenfassung und Überleitung

Das *Human Resources Performance* Modell differenziert in der Analyse und Planungsphase einer HF/HR/T-Maßnahme zwischen verschiedenen Betrachtungsebenen in der Steuerung und Evaluation. Diese Ebenen können auch als einander ergänzende Bewertungsebenen von HF/HR/T-Maßnahmen verstanden werden:

- (1) Organisationsebene: die Bewertung von HF/HR/T-Maßnahmen als Beitrag zur Umsetzung der Unternehmensstrategie (Schritt 1)
- (2) Prozessebene: die Bewertung von HF/HR/T-Maßnahmen als Unterstützung der Wertschöpfungs- und Leistungsprozesse (Schritt 2)
- (3) Funktions- oder Aufgaben-/Verhaltensebene: die Bewertung von HF/HR/T-Maßnahmen als Hilfe zur individuellen Aufgabenerfüllung und Förderung von anforderungsgerechter Qualifikation (Schritt 3)

Auf jeder dieser Ebenen kann der Nutzen einer HF/HR/T-Maßnahme mit Hilfe des HPM nachgewiesen werden. Für die Organisation ist die HF/HR/T-Maßnahme dann eine lohnende Investition, wenn sich die Verbesserung der individuellen Aufgabenerfüllung der geförderten Mitarbeiter in *Benefits* auf der Prozess- und Organisationsebene niederschlagen. Wie dies geschehen kann, wird im Folgenden in drei empirischen Evaluationen gezeigt.

Mit den drei empirischen Studien soll das HPM erfolgreich auf seine Anwendbarkeit und Nützlichkeit getestet werden. Zusammengefasst sind die wesentlichen, zu prüfenden Inhalte des HPM:

1. Die grundlegende Differenzierung der Wirtschaftlichkeitsbetrachtung auf den drei Ebenen „Funktion“, „Prozess“ und „Organisation“, die aufeinander aufbauen.
2. Die Modellierung des Evaluationsprozesses in Form von sieben Arbeitsschritten von der Analyse, über die Entscheidung/Planung bis zur Wirtschaftlichkeitsbewertung.
3. Die Trennung der Bewertung einer Personalmaßnahme in eine Wirkungseinschätzung und eine Bedeutungseinschätzung.

2 Empirische Untersuchungen

2.1 Gesamtkonzeption der drei empirischen Untersuchungen

Die drei empirischen Untersuchungen waren hinsichtlich ihrer Evaluationsobjekte, Fragestellungen, Quantität und Qualität der erhobenen Daten verschieden. Damit wurden eine breite Anwendung des HPM (Kapitel 1.4) und ein großes Erfahrungsspektrum hinsichtlich der ökonomischen Evaluation von HF/HR/T-Maßnahmen angestrebt. Obwohl die drei empirischen Untersuchungen das Kernstück dieses Berichtes sind und entsprechend viel Raum einnehmen – insbesondere Kapitel 2.2 – sind sie in der Gesamtkonzeption des Projektes nur Mittel zum Zweck, nämlich der Entwicklung und Verbreitung von Methoden zur ökonomischen Evaluation von HF/HR/T-Maßnahmen.

Zunächst wurde die in Kapitel 2.2.2 beschriebene Machbarkeitsstudie zur ökonomischen Evaluation des HF-Programms CISM ausgebaut. Damit zielte die Evaluationsfragestellung auf eine einzelne, seit einigen Jahren bestehende Maßnahme. Die Zielgruppen für die Evaluationsergebnisse waren alle Ebenen des *Managements* (*Top*, Bereich, Betrieb), die CISM-Fachverantwortlichen in der Unternehmenszentrale und den einzelnen Niederlassungen, aber auch die Mitarbeiter. Denkbar wäre auch, dass im Rahmen der anstehenden Kapitalprivatisierung von Flugsicherungsdiensten die Evaluationsergebnisse von Investoren positiv zur Kenntnis genommen werden. Damit sind in dieser Evaluationsstudie alle in Kapitel 1.3.1 identifizierten *Stakeholdergruppen* angesprochen.

Die CISM-Studie sollte vor allem zeigen, ob und wie die Verwendung der drei Analyseebenen des HPM (Organisationsebene, Prozessebene, Funktionsebene) für die ökonomische Evaluation hilfreich ist. Da CISM bereits seit Jahren eingesetzt wird, reduzierte sich die Anwendung des HPM allerdings auf die Bewertungsphase. Dennoch sollte mit dem 3-Ebenen-Modell das Herzstück des HPM getestet und weiterentwickelt werden. Insbesondere die damit verbundene umfassende und interdisziplinäre Datenerhebung und –auswertung sowie Beteiligung verschiedener Fachgruppen in der Organisation (Controlling, Personalmanagement, *Performance-Management*, Fachverantwortliche für CISM, *Management*) galt es, zu erproben.

Die zweite empirische Untersuchung (Kapitel 2.3) bezog sich nicht auf eine einzelne Maßnahme, sondern auf ein Maßnahmenbündel zur Personalentwicklung in einem *Call Center* der öffentlichen Verwaltung. Die besondere Herausforderung hier war die Isolierung der Wirkungen der einzelnen Maßnahmen und die Untersuchung ihres Zusammenspiels. Es wurden ausschließlich Führungskräfte und Mitarbeiter zum Nutzen der einzelnen Maßnahmen und des Programms insgesamt befragt. Organisations- oder Kostendaten wurden nicht zur Verfügung gestellt. Die Evaluation sollte vor allem die Arbeitsschritte der Funktions- und Wirkungsanalyse im HPM konkretisieren (Abbildung 1). Zielgruppe (Kapitel 1.3.1) der Evaluationsergebnisse sind in diesem Falle überwiegend die operativen Führungskräfte und der Bereichsleiter.

Die dritte empirische Untersuchung (Kapitel 2.4) hatte wiederum einzelne, in diesem Falle Trainingsmaßnahmen, als Evaluationsobjekte, nämlich *E-learning* Module, die in der Fluglotsenausbildung eingesetzt werden. Die besondere Anforderung an diese Studie war es, mit so wenig Aufwand wie möglich eine Investitionsentscheidung bzgl.

der zukünftigen Nutzung von *E-learning* herbeizuführen. Es wurden Auszubildende, Ausbilder und *Akademiemanager* befragt sowie die Kosten für die Entwicklung der Module erhoben. Leider wurden die Fragebogen nicht vollständig ausgefüllt, so dass Wirkungen des *E-learning*s auf Fähigkeiten/Fertigkeiten, Arbeitsverhalten und Arbeitsergebnisse der Mitarbeiter nicht dargestellt werden konnten. Die ökonomische Evaluation erfolgte ausschließlich über die durch das *E-learning* ersetzten traditionellen Lernformen. Die Auswertung in Kapitel 2.4 beschränkt sich auf diese Kostenreduktion durch ersetzten Klassen- und Simulatorunterricht und setzt diese ins Verhältnis zu den Entwicklungskosten des jeweiligen *E-learning* Moduls. Ziel war es, das Handlungsfeld „Ressourcenkontrolle“ im HPM (Abbildung 1) zu operationalisieren. An diesen Evaluationsergebnissen dürften insbesondere die Ausbilder, *Akademiemanager* aber auch externe Entwickler von *E-learning* interessiert sein (Kapitel 1.3.1).

2.2 Critical Incident Stress Management CISM

Diese erste Evaluationsstudie, die das *Critical Incident Stress Management* CISM der deutschen Flugsicherung (DFS) als Zielprogramm untersuchte, wird aus verschiedenen Gründen als *Benchmark* für die Evaluationspraxis betrachtet:

1. Es konnte eine Machbarkeitsstudie (Kapitel 2.2.2) durchgeführt werden, die die prinzipielle Möglichkeit und Sinnhaftigkeit einer ökonomischen Evaluation einer HF/HR/T-Maßnahme nachwies.
2. Im Verlauf der Machbarkeits- und der Hauptstudie erfolgte eine gründliche Vernetzung mit Fach- und Führungskräften der DFS. Daten und Erfahrungen aus verschiedensten *Management*bereichen (Sicherheits-, Qualitäts-, Personal-, Bereichs- und *Top-Management*) ergänzten die von uns durchgeführten Befragungen von Fluglotsen. Dadurch konnte die strategische und ökonomische Bedeutung der HF-Maßnahme klar, reliabel und valide herausgestellt werden.
3. Aus wissenschaftlicher Sicht setzte die Untersuchung wichtige Vorgaben um, die häufig in der Unternehmenspraxis vernachlässigt werden: Subjektive Aussagen von Lotsen und die unter 2. beschriebenen objektiven Daten untermauerten sich gegenseitig (*concurrent validity*). Eine weitere Validierung erfolgte durch quasi-objektive Daten, nämlich den Beobachtungen (Fremdeinschätzungen) von Führungskräften und CISM-Beratern. Die Veränderungen durch die HF-Maßnahme konnten in Relation zu einem Ausgangs- oder Normalwert bestimmt werden.

Im den folgenden Kapiteln 2.2.1 und 2.2.2 werden die Flugsicherung und die Machbarkeitsstudie kurz beschrieben. Dann erfolgt die Darstellung der Hauptstudie mit ihren Zielen, Methoden, Ergebnissen sowie deren Bedeutung für die ökonomische Evaluation von HF/HR/T-Maßnahmen im Allgemeinen.

2.2.1 Flugsicherung

Die Dienstleistung Flugsicherung ist für die Untersuchung der Steuerung und Evaluation von Personalmaßnahmen hervorragend geeignet, weil sie im Wesentlichen von Menschen, den Fluglotsen, vollzogen wird. Sie sorgen jeweils zu zweit (Vieraugenprinzip) für ausreichende Sicherheitsabstände (Staffelungen) zwischen den Flug-

zeugen in dem Luftraum, für den sie verantwortlich sind (Sektor). Auf der Grundlage des Flugplans, aus dem die zeitliche Folge und die Route der Flugzeuge hervorgehen, plant der koordinierende Fluglotse den Verkehrsfluss. Der exekutive Fluglotse hält den Funkkontakt zu den Piloten und gibt ihnen Anweisungen. Zwar wird ihre Arbeit von Technologie unterstützt, doch wesentliche Veränderungen dieser Technologie wurden erst in den letzten Jahren entwickelt und können entsprechen erst in naher Zukunft helfen, das steigende Flugverkehrsaufkommen zu bewältigen (Vogt, Adolph et al., 2002, Vogt, Hagemann & Kastner, im Druck). Damit repräsentiert die Fluglotsentätigkeit in besonderer Weise die Zukunftstrends unserer Arbeitswelt:

- Moderne Technologien (z.B. Datennetze zwischen Bodenstationen und Flugzeugen)
- Intensivierung der Arbeit (steigendes Flugverkehrsaufkommen)
- Zunehmende Komplexität (z.B. Vernetzung von Flugrouten, Anzahl und Verschiedenartigkeit der Flugzeuge)
- Steigende Dynamik (z.B. Erneuerung der Technologie und Organisation wie im *Single European Sky*)
- Internationalisierung (z.B. Arbeitssprache Englisch)
- Wissens- und fähigkeitsbasierte Dienstleistungsarbeit

Bei der Arbeit der Fluglotsen können kritische Ereignisse auftreten. Dies sind i.d.R. kleinere Unterschreitungen der Sicherheitsabstände (Staffelungsunterschreitungen), die von Passagieren und Piloten oft gar nicht wahrgenommen werden und für sie ohne Folgen sind. Tritt eine Staffelungsunterschreitung jedoch unvorhergesehen auf, wird sie von dem verantwortlichen Fluglotsen im Regelfall als Versagen wahrgenommen. Dies bedroht sein professionelles Selbstverständnis und kann *critical incident stress reactions* auslösen (Leonhardt & Vogt, im Druck). Um die Folgen solcher Stressreaktionen zu mildern, führte die DFS *Critical Incident Stress Management* (CISM) ein. Dabei handelt es sich um Methoden der Krisenintervention und psychischen ersten Hilfe, die von Kollegen für Kollegen angeboten werden (*Peer Model of Crisis Intervention*). Ein betroffener Fluglotse kann (auf freiwilliger Basis) diese kollegialen Berater nach einem kritischen Ereignis konsultieren. In CISM-Gesprächen soll dann eine schnellere Bewältigung des Ereignisses und eine langfristige Traumatisierung durch das Ereignis vermieden werden. Ziel von CISM ist sowohl die psychische Gesundung als auch die schnelle Wiederherstellung der Arbeitsfähigkeit. Derzeit sind 80 Fluglotsen der DFS in CISM ausgebildet. Die kollegialen Berater (*Peers*) verfügen über eine international anerkannte Ausbildung in CISM, die auch eine Schulung zum Erkennen von Störungen beinhaltet, die weiterer medizinischer oder psychologischer Hilfe bedürfen (*Referral*).

Das CISM-Programm der DFS war das HF-Zielprogramm im Rahmen der hier berichteten empirischen Untersuchungen. Eine Machbarkeitsstudie (Kapitel 2.2.2) sollte zunächst klären, ob die Anwendung ökonomischer Evaluationsmethoden auf ein solches Programm möglich und sinnvoll ist.

2.2.2 Zusammenfassung der Machbarkeitsstudie

Die Machbarkeitsstudie wurde 2003 durchgeführt, als das HPM (Kapitel 1.4) noch nicht entwickelt war. In Ermangelung eines integrativen Evaluationsmodells, das verhaltens- und wirtschaftswissenschaftliche Ansätze verbindet, wurde eine Kombination solcher Verfahren gewählt. Aus den Verhaltenswissenschaften wurde das Modell der Trainingsevaluation (Kirkpatrick, 1994) als theoretische Grundlage gewählt. Dabei handelt es sich um eine Taxonomie zur Erfassung der Trainingseffektivität, die in vielen Evaluationsstudien verwandt wurde. Kirkpatrick führte die Unterscheidung von vier Ebenen von Trainingseffekten ein:

1. Reaktionsebene: Sie beschreibt die Zufriedenheit der Teilnehmer mit der Trainingsmaßnahmen (Inhalte, Methode, Rahmenbedingungen, Trainerverhalten).
2. Lernebene: Auf dieser Ebene werden die Lerneffekte (Wissen, Fertigkeiten, Einstellungen) des Trainings erfasst unabhängig davon, inwieweit sie im praktisch anwendbar sind oder umgesetzt werden.
3. Verhaltensebene: Die dritte Ebene beinhaltet die Verhaltensänderungen bei der Arbeit, also den Lerntransfer und die Umsetzung der Lerneffekte durch die Teilnehmer am Arbeitsplatz. Hier werden arbeitsrelevante Verhaltensmerkmale eingeschätzt.
4. Ergebnisebene: Abschließend werden die (für die Organisation relevanten) Ergebniseffekte der Verhaltensänderungen gemessen: Produktqualität und Effizienz der Leistungserbringung, Kundenzufriedenheit und *Image*faktoren, Personal-faktoren (Mitarbeiter-zufriedenheit, Fluktuation, Abwesenheit) oder Kosteneinsparungen.

Die Idee dieses Ansatzes ist, dass die vier Ebenen aufeinander aufbauen und jede Ebene durch die vorangehende Ebene kausal beeinflusst wird. Das vier Ebenen-Modell von Kirkpatrick wurde von Phillips (1996) um eine 5. Ebene erweitert: die Bewertung der Rentabilität der Trainingsmaßnahme (*Return of Investment: ROI*). Dabei werden die Ergebnisgrößen in monetäre Einheiten übersetzt, die Vollkosten der Trainingsmaßnahme berechnet und daraus der ROI errechnet: $ROI = (\text{Gewinn} - \text{Kosten}) / \text{Kosten}$.

Die 5. Ebene der Trainingsevaluation ist das Herzstück der Nutzenanalyse (*Utility Analysis*, Boudreau & Ramstad, 2003) im Allgemeinen. Mit ihr wird die Wirtschaftlichkeit einer Maßnahme bewertet, indem die 5. Ebene, also die Ebene der ökonomischen Effekte, fokussiert wird. Ökonomische Bewertungsgrößen sind dabei der (zusätzliche) Ertrag einer Maßnahme (monetäre Leistungsverbesserung abzüglich der Kosten) oder die Rentabilität (Ertrag im Verhältnis zu den aufgewendeten Kosten).

Bezogen auf die Machbarkeitsstudie wurden diese Ansätze miteinander kombiniert. Ausgangspunkt war das CISM-Training der *Peers*: Diese wurden befragt hinsichtlich ihrer Reaktion auf die Trainings, der Lerneffekte und inwiefern das Gelernte in CISM-Gesprächen angewendet werden kann. Die *Peers* wurden also bezüglich der ersten drei Ebenen befragt. Die Fluglotsen, die von einem *Peer* beraten werden, sollen vom verbesserten Beratungsverhalten der *Peers* profitieren. Daher wurden sie gefragt,

wie ihre Reaktion auf die CISM-Konsultation war, ob und was sie dadurch gelernt haben, ob und was sich in ihrem Arbeitsverhalten geändert hat. Außer den ersten drei Ebenen wurden bei den Fluglotsen auch die Ebenen 4 und 5 erfasst, nämlich ob sich in der Folge eines kritischen Ereignisses bzw. einer CISM-Konsultation Arbeitsergebnisse und *Performance* auf Funktionsebene verändert haben und welche Bedeutung dies ggf. für die Organisations-*Performance* hatte.

Insgesamt nahmen 47 Fluglotsen und 13 *Peers* an der Untersuchung teil. Fragen zur Implementierung und den Kosten des Programms beantwortete der Fachverantwortliche. Abbildung 3 gibt eine Übersicht zu den Fragenbereichen auf der Grundlage des oben beschriebenen Evaluationsmodells.

Modell und Fragen an **Manager, Peers, Lotsen**

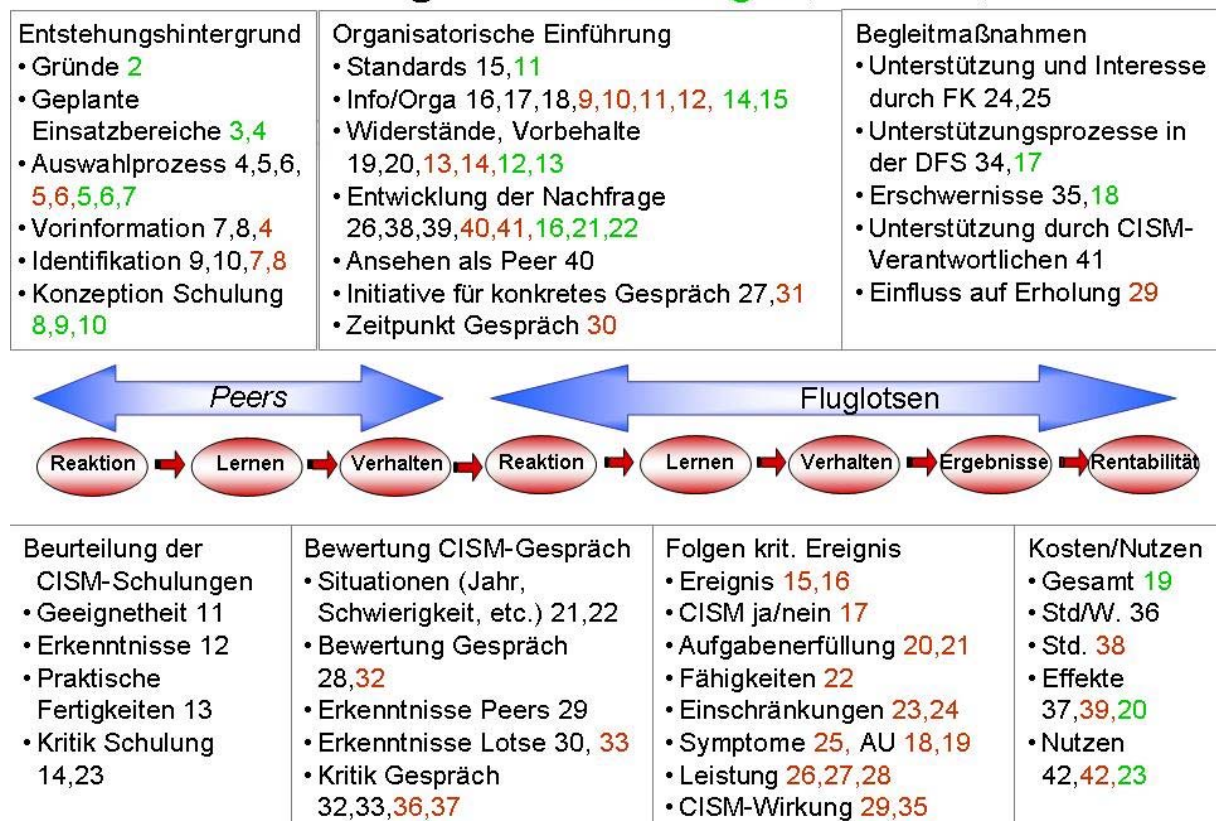


Abbildung 3: Evaluationsmodell und Fragenbereiche der Machbarkeitsstudie (die Zahlen verweisen auf die Frage- und Antwortnummern im Fragebogen bzw. Interview)

Die wesentlichen Ergebnisse der Machbarkeitsstudie waren wie folgt:

1. Die Fluglotsen berichteten von verschiedensten Ereignissen, in deren Folge sie einen *Peer* konsultierten. Darunter waren nicht nur Staffe-lungsunterschreitungen, sondern auch die Betroffenheit durch die Flug-zeugkollision über dem Bodensee im Jahr 2001 und private Belastun-gen.
2. Fluglotsen, die nach einem solchen Ereignis CISM in Anspruch nah-men, gaben im Durchschnitt an, dass die CISM-Gespräche an ihrer Er-holung einen Anteil von mit 36% hatten.

3. Bezogen auf die Gesamtdauer ihrer Arbeitsunfähigkeit bedeuteten diese 36%, dass Fluglotsen nach Staffelungsunterschreitungen im Durchschnitt einen Tag schneller zum Arbeitsplatz zurückkehrten, als sie dies eigenen Angaben zu Folge ohne CISM getan hätten. Nach der Flugzeugkollision über dem Bodensee betrug diese Zahl drei Arbeitstage.
4. In der Untersuchungsstichprobe ergaben sich auf diese Weise 47 gewonnene Arbeitstage, deren Wert beziffert werden konnte. Dieser Wert durch CISM gewonnener Leistungszeit war das Zweieinhalbfache der Kosten, die das Programm – bezogen auf die Stichprobe – in den Jahren seines Bestehens verursacht hat.⁵
5. Der ROI betrug damit 250%.
6. Darüber hinaus ergab die Machbarkeitsstudie Hinweise auf weitere, intangible Effekte, wie z.B. eine Verbesserung der Vertrauens-Fehler-Lernkultur (Kastner, 2006). Auch Fluglotsen, die nach einem kritischen Ereignis kein CISM in Anspruch nahmen, berichteten von solchen Nutzeneffekten. Sie lobten die positiven Wirkungen des Programms auf den Umgang mit Fehlern in ihrem Unternehmen und schätzten CISM als *Back-up* für den Fall, dass sie es in Zukunft benötigen würden.

Die Machbarkeitsstudie hat die prinzipielle Anwendbarkeit ökonomischer Evaluationsmethoden auf HF-Programme nachgewiesen. Sie ergab darüber hinaus Hinweise auf vielfältige intangible Nutzeneffekte. Beides sollte in der Hauptstudie vertieft untersucht werden.

2.2.3 Ziele der Hauptstudie

Die Machbarkeitsstudie (Kapitel 2.2.2) lieferte deutliche Hinweise für die Wirtschaftlichkeit von CISM. Allerdings hatte diese Studie nicht zum Ziel, den Nachweis der Wirtschaftlichkeit zu erbringen; sie sollte lediglich prüfen, ob und mit welchen Methoden man die Wirtschaftlichkeit eines Instrumentes wie CISM überhaupt nachweisen kann. Darauf aufbauend sollen nun wissenschaftlich zuverlässige und gültige Beweise für die Auswirkungen von CISM auf die Produktivität und Kapazität der Flugsicherung geliefert werden sowie die Hinweise auf Prozess- und Kulturverbesserungen verfolgt werden. Bezogen auf die Gesamtkonzeption der empirischen Untersuchungen in dem vorliegenden Bericht sollte diese *Benchmark*-Evaluation hinsichtlich ihres HPM-Einsatzes sowie der Quantität und Qualität der Datenerhebung und Auswertung zum *Benchmark* für die ökonomische Evaluation von HF/HR/T-Maßnahmen werden (Kapitel 2.1).

Die Machbarkeitsstudie (Kapitel 2.2.2) soll in drei Punkten durch die Hauptstudie ergänzt werden:

⁵ Bereits in der Machbarkeitsstudie wurde ein Prototyp der *Performance*-Kurve (Kapitel 2.2.5.3) verwendet, auf dessen Grundlage diese Berechnungen vollzogen wurden.

1. Schaffung einer soliden Basis für eine Wirkungs- und Wirtschaftlichkeitsaussage:

- Vergrößerung der Stichprobe und ihrer Repräsentativität
- Eingrenzung des Zeitraums berichteter kritischer Ereignisse auf zwei Jahre; ein kritisches Ereignis ist z.B. die Unterschreitung des Sicherheitsabstandes zwischen zwei Flugzeugen (Staffelungsunterschreitung)
- Eingrenzung der kritischen Ereignisse auf Unterschreitungen der Sicherheitsabstände von Flugzeugen (Staffelungsunterschreitungen)
- Erfassung aktueller kritischer Ereignisse im Untersuchungszeitraum

Die Angaben zu den kritischen Ereignissen, der Wirkung von CISM und der Erholung waren in der Machbarkeitsstudie (Kapitel 2.2.2) zum Teil wenig präzise und sehr heterogen. Dies lag daran, dass viele der in der Machbarkeitsstudie beschriebenen Ereignisse schon weit mehr als zwei Jahre zurücklagen und teilweise in den privaten Bereich fielen (z.B. Tod des Ehepartners). Für eine zeitnahe und homogene Erhebung war die damalige Stichprobe zu klein.

2. Klärung der Wirkungskette auf der funktionalen Ebene (Fluglotse):

Die Machbarkeitsstudie (Kapitel 2.2.2) ermöglichte die Ableitung von Hypothesen zur Wirkungsweise von kritischen Ereignissen und des CISM-Programms auf den individuellen Fluglotsen und die Unternehmenskultur der DFS. Diese Hypothesen sollten überprüft werden.

3. Einbeziehung der Prozessebene und der Organisationsebene aus dem *Human Resources Performance Modells* (HPM):

In der Machbarkeitsstudie (Kapitel 2.2.2) wurden die monetären Effekte aus der beschleunigten Rückkehr der Fluglotsen nach kritischen Ereignissen berechnet. Entscheidend für die realen, wirtschaftlichen Effekte ist aber die Auswirkung von kritischen Ereignissen und CISM auf den Prozess der Flugsicherung. Die Wirtschaftlichkeit der Investitionsentscheidung CISM hängt dann zusätzlich von der Bedeutung der Effizienz und Effektivität des Flugsicherungsprozesses für den Flugsicherer ab, z.B. der Bewertung von Kosten- und Qualitätseffekten in der Flugsicherung mit Blick auf die Fluggesellschaften als Kunden und andere europäische Flugsicherungen als potentielle Wettbewerber (*Single European Sky*). Die Hauptstudie sollte untersuchen, wie sich durch kritische Ereignisse ausgelöste Leistungsveränderungen und Erholungseffekte von CISM auf zentrale Parameter im Flugsicherungsprozess auswirken und wie der Prozess nach kritischen Ereignissen gesteuert wird. Die Auswirkungen werden im Hinblick auf die strategischen Ziele und übergeordneten Kosten- und Nutzenkennzahlen betrachtet und bewertet.

2.2.4 Der Evaluationsprozess in der Hauptstudie

Um die drei Ebenen (funktionale Ebene, Prozessebene, Organisationsebene) und deren Vernetzung in der Hauptstudie zu erfassen, galt es, neben den Fluglotsen auch die Prozessexperten und die obere Führungsebene in die Untersuchung einzubeziehen. Die unterschiedlichen Auskunftgeber und zu erfassenden Informationsbereiche zeigt Abbildung 4.

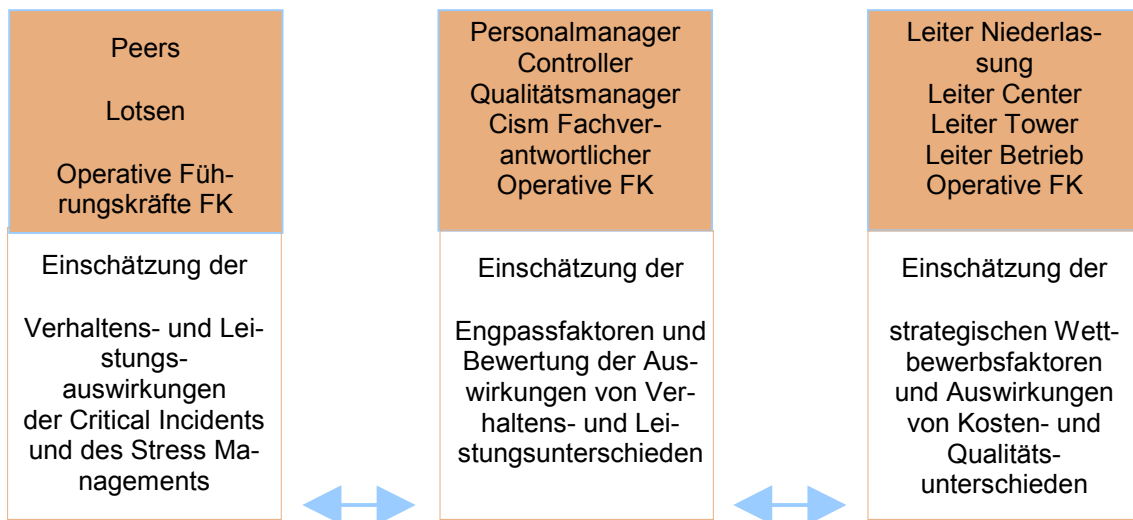


Abbildung 4: Die verschiedenen Auskunftgeber und Informationsbereiche in der Evaluation des *Critical Incident Stress Managements* der Flugsicherung

Der Evaluationsprozess wurde in drei Schritte eingeteilt:

1. Modellierung der Zusammenhänge von Verhalten und Wirtschaftlichkeit

Zunächst wurden in *Workshops* und Interviews die Fach- und Führungskräfte zu den Kriterien, Kennzahlen und Interdependenzen befragt, die in der Untersuchung zu berücksichtigen sind. Ziel war es, eine Datenarchitektur für die Befragung und Auswertung zu erhalten.

2. Befragung zu Verhaltens- und Leistungsauswirkungen von kritischen Ereignissen

Anhand der Datenarchitektur wurden schriftliche Befragungen von Fluglotsen und *Peers*⁶ konzipiert und anschließend durchgeführt und statistisch ausgewertet.

3. Ökonomische Bewertung der Wirkungen auf Prozess- und Organisationsebene

Die Ergebnisse der Befragung wurden anschließend wieder mit Fach- und Führungskräften analysiert und ökonomisch bewertet.

⁶ *Peers* oder kollegiale Berater werden die Fluglotsen genannt, die einen Kollegen nach einem kritischen Vorfall betreuen; sie wurden (aus)gewählt und entsprechend den Vorgaben der *International Critical Incident Stress Foundation* für diese Betreuungsaufgabe weitergebildet.

2.2.5 Modellierung der Zusammenhänge von Verhalten und Wirtschaftlichkeit in der Hauptstudie und Konzeption der Erhebungsinstrumente

Ausgehend von der Machbarkeitsstudie (Kapitel 2.2.2) wurde zunächst ein Basismodell für die Datenerhebung erstellt. Es zeigt auf, dass die Untersuchung zwei verschiedene Datenbereiche erfassen und miteinander verbinden will, die traditionell von verschiedenen Disziplinen erforscht werden (s. Abbildung 5):

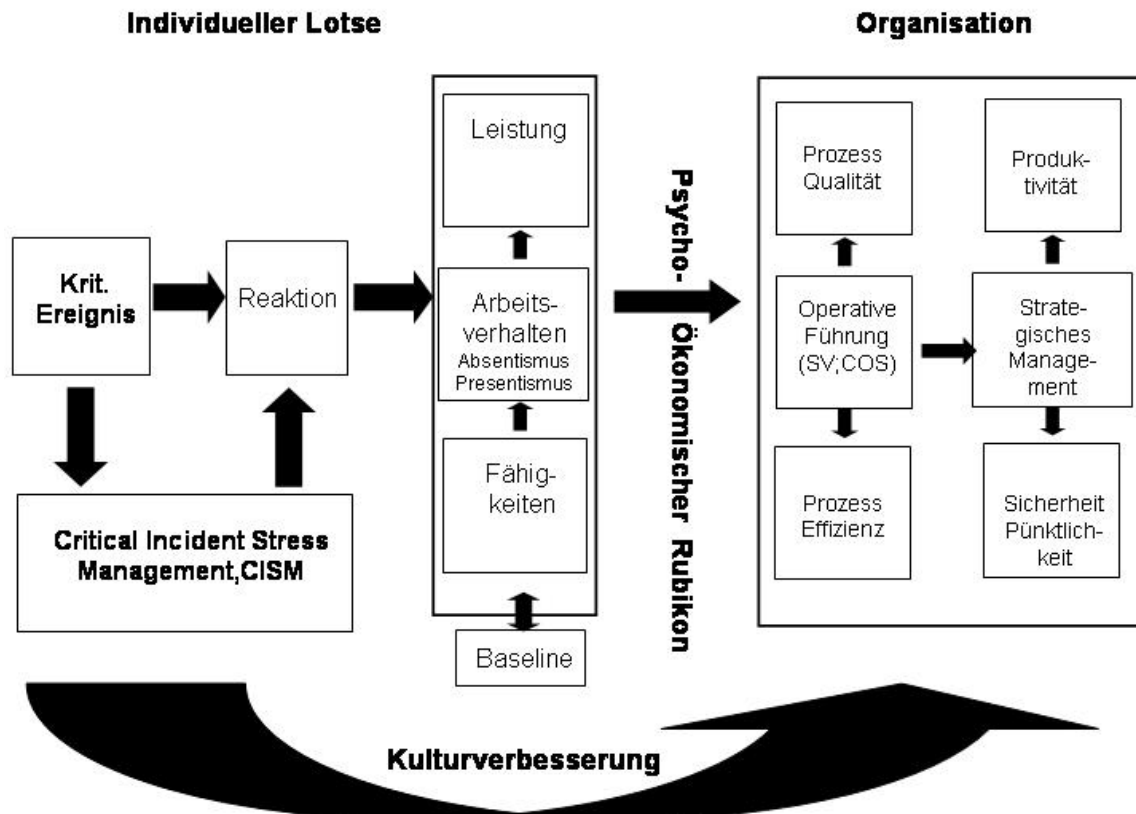


Abbildung 5: Die *Benchmark-Evaluation* verbindet die Individual- mit der Organisationsebene und überschreitet so den psycho-ökonomischen Rubikon (SV: *Supervisor*, COS: *Chief of Section*)

Der linke Bereich der Abbildung 5 stellt die individuelle Verarbeitung eines kritischen Ereignisses dar, wie sie in den Verhaltenswissenschaften untersucht wird.

Der rechte Bereich der Abbildung 5 verweist auf die ökonomischen Faktoren und die Rolle des Managements, wie sie in den Wirtschaftswissenschaften betrachtet werden.

Die Verbindung von Verhalten und Wirtschaftlichkeit ist bislang weitgehend ignoriert worden. Arbeits- und Organisationspsychologie auf der einen Seite und Wirtschaftswissenschaften auf der anderen Seite sind überwiegend getrennte Disziplinen. Dieser „psycho-ökonomische Rubikon“ soll in der hier berichteten Hauptstudie überschritten werden. Der Name soll auf die Tatsache hinweisen, dass im ökonomischen Management und *Controlling* die Verhaltensvariabilität als Planungsgröße bislang unbeachtet ist. Personal- und Prozesskosten werden mit *Head Counts* geplant, bei denen jeder Mitarbeiter unabhängig von Qualifikation, Gesundheit, Arbeitseinstellung oder Motivation gleich viel zählt. In verhaltenswissenschaftlichen Untersuchungen

werden auf der anderen Seite in der Regel Verhaltens- und Leistungsniveaus als Kriterium für wirtschaftliche Effekte verwandt. Dass Verhaltens- und Leistungsunterschiede ökonomisch bedeutsam sind, wird oft angenommen, aber nicht explizit untersucht oder es werden vereinfachende Annahmen über deren Bedeutung gemacht wie z.B. in der Nutzwertanalyse. Der Zusammenhang zwischen Verhalten und Wertschöpfungsprozess ist hier eine *Black Box*.

2.2.5.1 Workshops mit den Fachkräften

In einem ersten Schritt sollte das psycho-ökonomische Rubikon-Modell für den Bereich der Flugsicherung operationalisiert werden. Um der Vielschichtigkeit des Vorhabens gerecht zu werden, wurden Experten der DFS aus den Bereichen „Personal“, „*Performance* und *Controlling*“ sowie der CISM-Fachverantwortliche in diesen Operationalisierungsschritt eingebunden. Es sollte geklärt werden

1. welches die zentralen Herausforderungen für die DFS aktuell sind und in den nächsten Jahren sein werden (Markt, Wettbewerb und *Change* Prozesse),
2. welche Zielsetzungen daraus für die Unternehmung abgeleitet worden sind und verfolgt werden (strategische Ziele),
3. wie diese Ziele erreicht werden sollen (strategische Handlungsfelder),
4. wie diese Ziele gemessen werden (strategische Kennzahlen),
5. wer an welchen Zielen gemessen wird (Führungsverantwortlichkeiten),
6. was dies für das Personalmanagement bedeutet (Personalstrategie),
7. welche Anforderungen an den Prozess der Flugsicherung und den einzelnen Fluglotsen sich daraus ergeben (Prozessmerkmale und Lotsenanforderungsprofil),
8. wie die Leistung des einzelnen Fluglotsen und die Kosten und Qualität des Flugsicherungsprozesses gemessen werden (*Performance*-Indikatoren),
9. welche aktuellen Maßnahmen zur Leistungs- und Prozessoptimierung geplant und umgesetzt werden (Prozessmanagement), und
10. welche Auswirkungen individuelle Verhaltensunterschiede und Variabilitäten im Verhalten der Fluglotsen auf die Prozessstabilität, die *Gesamt-Performance* und die Personalplanung haben und welche Bedeutung das individuelle Verhalten im Personal- und *Performance*-Management der DFS spielt (Verhaltensmanagement).

Aus den Ergebnissen der *Workshops* sollte eine Landkarte der für diese Untersuchung relevanten Parameter und Kennzahlen sowie deren Abhängigkeiten erstellt werden.

Die *Workshops* erwiesen sich als wahrer *think tank*, da die Experten mit großer Offenheit, Konstruktivität und Ideenreichtum die Fragen diskutierten. Es entstand ein tiefes Verständnis der Gesamtzusammenhänge zwischen Lotsenverhalten und Unternehmens*performance*, das in Form von drei Steuerungskreisläufen visualisiert wurde (s. Abbildung 6).

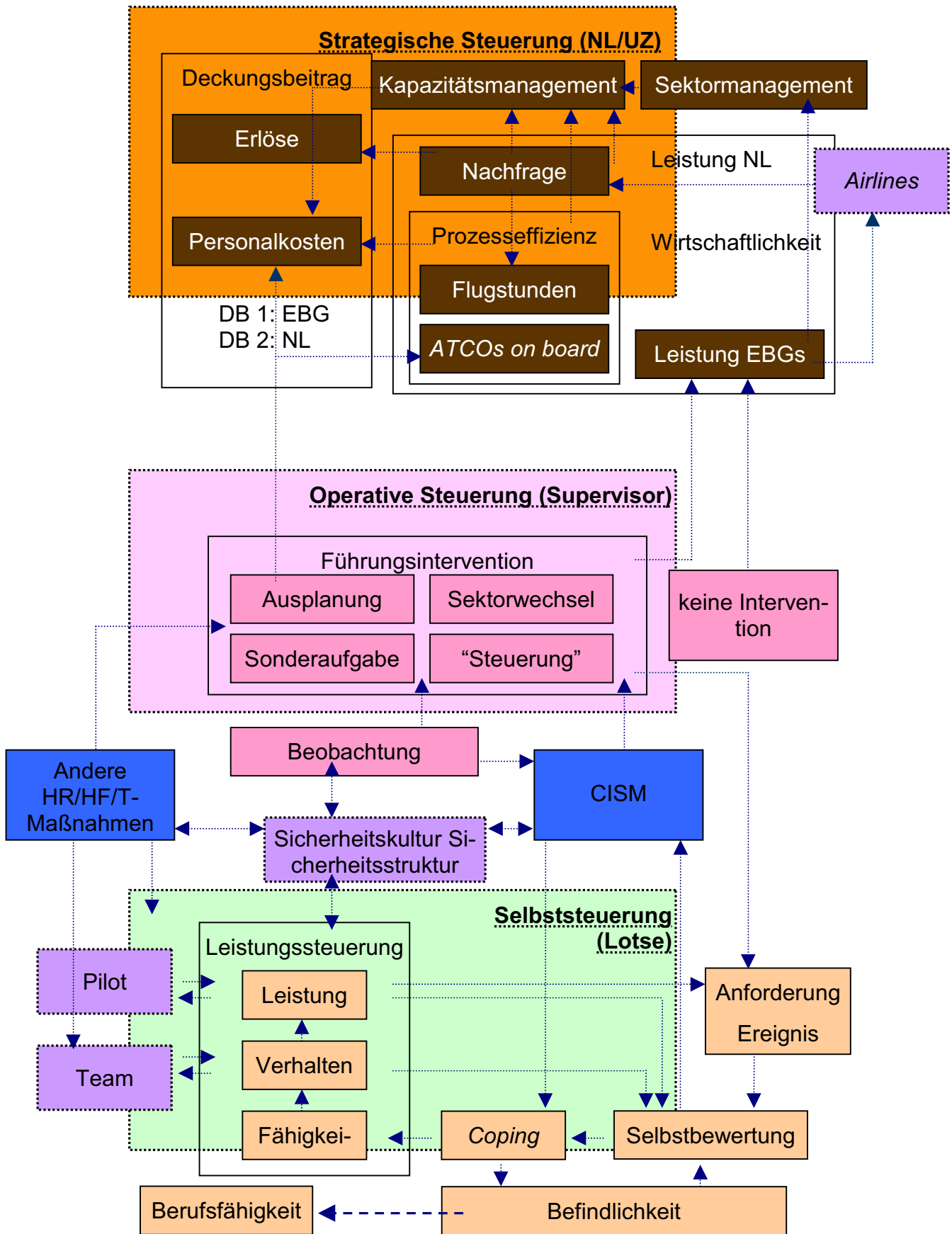


Abbildung 6: Steuerungskreisläufe im Flugsicherungsprozess (DB: Deckungsbeitrag, EBG: Einsatzberechtigungsguppe, NL: Niederlassung, UZ: Unternehmenszentrale)

Die dargestellten drei Steuerungskreisläufe sind (beginnend am unteren Bildrand)

- die Selbststeuerung des Fluglotsen
- die Prozesssteuerung durch die Führungskräfte vor Ort
- die zentrale, strategische Steuerung durch das TOP-Management

Die Steuerungskreisläufe entsprechen den drei Ebenen im HPM (Kapitel 1.4).

1. Der Steuerungskreislauf „Fluglotse“ stellt die Selbstregulation des Fluglotsen nach einem kritischen Ereignis dar. Dieser Kreislauf basiert auf den Erkenntnissen aus der Machbarkeitsstudie (Kapitel 2.2.2, siehe auch Vogt et al., 2004).
2. Der zweite Kreislauf betrifft die „operative Steuerung“. Sie sollte den Fluglotsen dort unterstützen, wo die Selbstregulation überfordert ist oder sich negativ auf die Zusammenarbeit im Arbeitsprozess und auf die Prozessziele auswirkt. Die beiden letztgenannten Aspekte betrachten das Lotsenteam (früher bezeichnet als „Wache“, heute „Einsatzberechtigungsgruppe“ EBG) als die nach dem individuellen Fluglotsen kleinste operativ zu steuernde Einheit. Die operative Führung (*Supervisor*) sollte dann eingreifen, wenn solche dysfunktionalen Muster erkennbar sind. Die Beobachtung und Bewertung der Regulation durch den *Supervisor* hängt dabei entscheidend von den internen Sicherheitsstandards (Sicherheitsstruktur, d.h. alle Sicherheitsinstrumente der DFS wie z.B. der vierteljährliche *Safety Letter* über kritische Ereignisse) und der gelebten Sicherheitskultur ab. Die Führungskraft handelt in diesem Kontext und der definierten Führungsrolle. Grundlegende Handlungsmöglichkeiten in der operativen Führung sind:
 - „Steuerung“ (im extremsten Fall wird der Flugsicherungssektor geschlossen, d.h. kein weiteres Flugzeug darf einfliegen)
 - „Ausplanen“ (Ablösung des Fluglotsen bzw. Beurlaubung)
 - „Sonderaufgabe⁷“ (Herausnahme des Fluglotsen aus der operativen Tätigkeit)
 - „Sektorwechsel“ (Einsatzwechsel auf einen anderen operativen Arbeitsplatz mit geringerer Arbeitslast)
 Die Eingriffsmöglichkeiten werden durch HR/HF/T-Maßnahmen erweitert. CISM ist eine solche Maßnahme, die der *Supervisor* vorschlagen kann.
3. Auf der obersten Ebene ist die strategische Steuerung angesiedelt. Dort werden die Daten aus dem Leistungsprozess erfasst. Wesentliche Leistungskenngrößen der Flugsicherung sind:
 - Flugsicherheit (z.B. Anzahl kritischer Ereignisse)

⁷ D.h. eine nicht-operative, aber dennoch indirekt wertschöpfende Tätigkeit wie z.B. die Entwicklung neuer An-/Abflugverfahren.

Produktivität (z.B. Anzahl kontrollierter Flugstunden pro Lotsenteam, das sind alle Fluglotsen, die eine bestimmte Gruppe Sektoren überwachen, auch Einsatzberechtigungsgruppe EBG genannt)

Effizienz (Flugstunden pro diensthabendem Fluglotsen)

Wirtschaftlichkeit (Erlöse - Personalkosten)

In Abhängigkeit von diesen Ergebnissen werden zentral Sektorgrößen und Personalkapazitäten geplant.

2.2.5.2 Interviews mit den Führungskräften

Da die schriftliche Lotsenbefragung in verschiedenen Niederlassungen der DFS durchgeführt werden sollten, wurden die leitenden Führungskräfte in diesen Niederlassungen über das Vorhaben informiert und zugleich zu ihrer Sicht der Bedeutung von CISM befragt. Hier sollte vor allem der größere Zusammenhang zwischen der verhaltensorientierten Maßnahme und den strategischen und wirtschaftlichen Zielen der Niederlassungen aufgezeigt und konkretisiert werden.

In jeder Niederlassung (NL) wurden 2-3 stündige Interviews geführt. Teilnehmer waren insgesamt 11 NL-Leiter, 39 operative Führungskräfte (*Supervisors SV* und *Chiefs of Sections COS*) aus den Niederlassungen (NL) sowie 28 *Peers*. Die *Peers* wurden bei dieser Gelegenheit vor allem auf ihre Aufgabe der Verteilung der Lotsenfragebogen vorbereitet. Die Leitfragen für die Gespräche mit den NL-Leitern und Führungskräften wurden in einem strukturierten Interviewleitfaden mit vier Themenschwerpunkten ausgearbeitet:

1. Strategie, Ziele und Veränderungsthemen der NL

- Was sind für Ihre Niederlassung die wichtigsten Ziele der nächsten Jahre?
- Was sind die zentralen Erfolgsmerkmale?
- Was sind die besonderen Herausforderungen für die Niederlassung?
- Was soll sich ändern? Woran müssen Sie „arbeiten“?

2. Produktivität, Effizienz und Qualität

- Was sind die wichtigsten Maßnahmen zur Steigerung der Qualität?
- Welches sind die wichtigsten Aktivitäten zur Optimierung und Effizienzsteigerung?
- Wie versuchen Sie in der Niederlassung die Produktivität zu beeinflussen?
- Welche Kennzahlen gibt es in der NL, um dies zu steuern?
- Wer kümmert sich darum?

3. Leistungsfähigkeit der Fluglotsen

- In welcher Form und mit welchen Maßnahmen (neben CISM) wird die Leistungsfähigkeit der Fluglotsen in Ihrer Niederlassung beachtet und gefördert?
- Welche Bedeutung hat dies für die NL?

- Welche Bedeutung spielt hier die individuelle Leistungsfähigkeit?
- Inwieweit beachten Sie Unterschiede in der Leistungsfähigkeit und der Leistung von Fluglotsen?
- Welche Auswirkungen hat es, wenn Fluglotsen in ihrer Leistungsfähigkeit am Arbeitsplatz (durch mangelnde Gesundheit, eingeschränkte psychische Belastbarkeit, Konzentrationsmängel) beeinträchtigt sind?
- Welche Bedeutung haben kritische Ereignisse im Flugverkehr für die NL?

4. Personalmanagement in der NL

- Welche Möglichkeiten haben Sie, Personalkosten zu steuern und einzusparen?
- Welche Bedeutung hat es dabei, die Leistungsfähigkeit der Fluglotsen (als Gruppe und/oder individuell) weiter zu entwickeln oder zu bewahren?
- Welche Maßnahmen gibt es dazu?
- Wie können Sie das Verhältnis von Kapazität und Personalkosten optimieren?
- Inwieweit wird bei der Einsatzplanung die individuelle und aktuelle Einsatzfähigkeit berücksichtigt?

Die Interviews brachten wichtige Erkenntnisse über die aktuellen Handlungsfelder in den Niederlassungen und verdeutlichten die Rahmenbedingungen unter denen CISM dort eingesetzt wird.

Handlungsfeld „Wirtschaftlichkeit“

Die Niederlassungen steuern ihre Arbeit zunehmend durch Kennzahlen und professionalisieren ihr *Controlling*. Dies umfasst die Aspekte *Safety*, Produktivität und Pünktlichkeit. Insbesondere im Bereich *Safety* werden die Instrumente zunehmend standardisiert. *Safety*-Kennzahlen werden den Fluglotsen im operativen Geschäft transparent gemacht. Um eine hohe Produktivität zu erreichen, achtet man auf einen optimalen Verkehrsfluss und eine Optimierung des Gesamtsystems im Sinne von Effizienz. Prozesseffizienz ist definiert als Anzahl von Flugstunden pro operativ tätigem Fluglotsen. Die Wirtschaftlichkeit ergibt sich dann durch eine hohe Prozesseffizienz verbunden mit einer hohen Lotsen*performance* bei gleichzeitig hoher Nachfrage. Die Nachfrage ist dabei der Faktor, der nur bedingt aktiv gesteuert werden kann, z.B. durch das Anbieten von Ausweichstrecken, die – vergleichbar einer Umleitungsempfehlung auf der Autobahn – zwar länger, aber weniger beflogen sind.

Hintergrund dieses Handlungsfeldes ist vor allem die Sicherung des eigenen Standortes im Rahmen der aktuellen und zukünftigen Konzentrationsprozesse. Sicherheit ist hier eine Bedingung *sine qua non*. Durch ein gutes Preis-Leistungs-Verhältnis soll darüber hinaus ein Wettbewerbsvorteil im europäischen Markt geschaffen werden. Dabei wird die Anpassungsfähigkeit der Organisation als wichtiger Erfolgsfaktor gesehen. Nur durch permanente Weiterentwicklung der Technik, der Arbeitsprozesse und Arbeitsstrukturen wird die zukünftige Wettbewerbsfähigkeit gewährleistet. Die Komplexität der Steuerungsmechanismen vor Ort (z.B. Sektormanagement) nimmt dabei zu. Um Märkte zu besetzen, werden auch Kooperationen mit anderen Organi-

sationen eingegangen (z.B. Flughäfen oder andere europäische Flugsicherungen, *Air Navigation Service Providers* ANSPs genannt).

Handlungsfeld Kapazität

Die Optimierung der Personalkapazität ist nicht nur ein strategisches Handlungsfeld in der Unternehmenszentrale sondern auch für die Niederlassungen. Es gilt, *Performance*-Kapazitäten an den wachsenden Luftverkehr anzupassen bei konstanten oder sogar reduzierten Personalressourcen. Die Niederlassungen werden dabei am so genannten Deckungsbeitrag II gemessen, der die Erlöse in das Verhältnis zu den Personalkosten setzt. Die Personalkosten können allerdings in den Niederlassungen nur gesenkt werden, wenn die Prozesseffizienz und Fluglotsen-*Performance* gesteigert wird und gleichzeitig mittel- und langfristig Mitarbeiterkapazitäten auch in den indirekten Bereichen z.B. Flugdatenbearbeitung abgebaut werden. Zu diesem Zweck werden ggf. Fluglotsen auch in nicht direkt wertschöpfenden Aufgaben eingesetzt (Sonderaufgaben, s.u.), um weniger qualifiziertes und weniger gebundenes Personal freisetzen zu können.

Handlungsfeld Fluglotsen-*Performance*

Die Erreichung der strategischen Ziele hängt in hohem Maße von der Leistungsfähigkeit der Fluglotsen ab. Hier gibt es eine Vielzahl von Unterstützungsmaßnahmen in den Bereichen HF/HR/T. Aus der Sicht der Niederlassungen sollten diese Maßnahmen noch stärker abgestimmt und integrativ gesteuert werden. Es wurde aber auch deutlich, dass in hohem Maße die Eigenverantwortung der Fluglotsen bei der Leistungssteuerung gefordert ist. Es gilt das Prinzip, dass jeder Fluglotse jederzeit zu 100% flexibel einsetzbar ist. Dabei wird davon ausgegangen, dass entweder der Fluglotse selbst seine eigenen Grenzen erkennt und diese äußert („Der Fluglotse wählt seinen Arbeitsplatz selbst oder äußert seine Tagesform“), oder dass eine Selbstregulation im Team erfolgt. Auch bei der Fitness und Stresstabilität wird zunehmend auf die Selbstverantwortung des Fluglotsen gesetzt. Die Selbsteinschätzungskompetenz des Fluglotsen, sein Verantwortungsbewusstsein und die Qualität der Teamarbeit werden somit zum kritischen Faktor insbesondere nach kritischen Ereignissen. Individuelle Leistungsschwankungen nach kritischen Ereignissen sollten durch das Team kompensiert werden. Im Personalmanagement und der Einsatzplanung wird der Faktor „Leistungs- und Verhaltensvariabilität“ jedoch nicht systematisch gemessen oder berücksichtigt. Gleichzeitig wird vermutet, dass eine persönliche Unterstützung im Team durch die Zusammenlegung von NL und dadurch wechselnde Teamkonstellationen erschwert wird und daher die institutionalisierte Unterstützung, wie durch CISM, an Bedeutung gewinnt.

Handlungsfeld CISM

Bei den Interviews wurde klar und deutlich angesprochen, dass CISM nicht als Maßnahme betrachtet werden sollte, die sich unter wirtschaftlichen Aspekten lohnt. Vielmehr wurde betont, dass CISM als wichtige Unterstützungsmaßnahme bei Fluglotsen und Führungskräften akzeptiert ist und für andere HF/HR/T-Maßnahmen Modellcharakter hat. Gleichwohl erfolgt die Umsetzung der CISM-Standards in den Niederlassungen unterschiedlich.

Die Interviews bestätigten somit einerseits die Ergebnisse der *Workshops*. Andererseits wurde auch deutlich, dass der Bezug zwischen HF/HR/T-Maßnahmen und den aktuellen und zukünftigen Anforderungen im europäischen Wettbewerb bisher wenig konstruiert wird.

2.2.5.3 Fragebogen für Fluglotsen

Im Anschluss an die *Workshops* und Interviews wurde die schriftliche Befragung konzipiert. Sie sollte vor allem dazu dienen, den ersten Steuerungskreislauf „Fluglotse“ zu quantifizieren und Bezüge zum zweiten Steuerungskreislauf „operative Steuerung“ aufzuzeigen. Eine objektive Leistungs- und Verhaltensmessung für Fluglotsen konnte nicht realisiert werden. Wie in der Machbarkeitsstudie (Kapitel 2.2.2) musste auf Selbst- und Fremdeinschätzungen von Fluglotsen und *Peers* zurückgegriffen werden. Allerdings konnte mit Hilfe einer stichprobenartigen Überprüfung der Verkehrssituationen im Zusammenhang mit kritischen Ereignissen, die routinemäßig in der *Human Error Analysis* (HERA) der DFS durchgeführt wird, festgestellt werden, dass die Fluglotsen ihre Arbeitsbelastung und –leistung sehr realistisch einschätzen. Um die Akzeptanz der Datenerhebung zu erhöhen, wurde zugleich auf eine Fremdeinschätzung durch die Führungskräfte verzichtet, da eine Mitarbeiterbeurteilung noch nicht in systematischer Form im Unternehmen etabliert war. Um die Reliabilität der Selbsteinschätzung zu überprüfen, wurde stattdessen eine Fremdeinschätzung durch den *Peer* beschlossen, der den betroffenen Fluglotsen nach dem kritischen Ereignis betreut hat.

Die Konzeption des Fragebogens anhand der beiden Steuerungskreisläufe (Fluglotse, operative Führung) führte zu folgender Fragebogenstruktur:

Abschnitt 1.: Erhebung der Basislinie an allen befragten Fluglotsen

Um den Einfluss des kritischen Ereignisses und von CISM auf die Lotsenleistung beurteilen zu können, sollten alle befragten Fluglotsen zunächst ein situationsunabhängiges, persönliches Stärken-Schwächen-Profil ihrer Fähigkeiten, ihrer Tätigkeitsausübung und ihrer Leistungsergebnisse erstellen. Für die Zusammenstellung der Items wurden vorhandene Anforderungsprofile von Fluglotsen bei der DFS (Deichert & Eißfeld, 1998; Fleishman & Reilly, 1992) und bei EUROCONTROL (2005) ausgewertet und in diese drei Bereiche gruppiert. Das Gesamtprofil erfasst 23 Fähigkeiten (z.B. Reaktionszeit), die für die Ausübung des Berufs notwendig sind, die 14 wesentlichen Tätigkeiten eines Fluglotsen (z.B. Verkehrsbild beobachten/bewerten), sowie die 5 zentralen Ergebnismerkmale (neben *Safety*, Kapazität und Flüssigkeit der Verkehrsabwicklung auch die Teamleistung und die Erfüllung von Kundenanforderungen).

Abschnitt 2.: Erhebung der Wirkung eines in der Vergangenheit erlebten kritischen Ereignisses

Im zweiten Abschnitt wurden die Fluglotsen gebeten, ein kritisches Ereignis aus den letzten zwei Jahren zu beschreiben. Weiter zurückliegende Ereignisse wurden ausgeschlossen, weil davon ausgegan-

gen wurde, dass die Beschreibung von jüngeren Ereignissen zuverlässiger ist. Die Beschreibungsmerkmale entsprechen den Variablen im ersten und zum Teil zweiten Steuerungskreislauf:

- (1) Beschreibung des Ereignisses und der unmittelbaren Auswirkungen auf die eigene Befindlichkeit (emotional, körperlich, mental, im Verhalten)
- (2) Beschreibung der unmittelbaren Reaktionen anderer und Konsequenzen des Ereignisses für Kollegen und Führungskräfte
- (3) Angabe der Dauer des gesamten Belastungszeitraums, d.h. Zeitraum bis zur vollen Wiederherstellung der eigenen Leistungsfähigkeit
- (4) Darstellung des Verlaufs der Erholung der eigenen Leistung im Zeitraum von acht Wochen (*Performance-Kurve*)
- (5) Darstellung der Veränderungen von Fähigkeiten, Tätigkeitsausübung und Leistungsergebnissen im Belastungszeitraum im Verhältnis zur Basislinie
- (6) Beschreibung der Häufigkeit von emotionalen und körperlichen Beeinträchtigungen im gesamten Belastungszeitraum
- (7) Beschreibung und Quantifizierung der Auswirkungen der Beeinträchtigungen auf den gemeinsamen Arbeitsprozess in sieben Kategorien (*Safety*, Arbeitsfehler, Kapazität, Verkehrsabwicklung, Zusammenarbeit im Sektor, sektorübergreifende Zusammenarbeit, Personaleinsatz)
- (8) Beschreibung der *Coping*strategien und des Einflusses von CISM auf die Strategien sowie der Bedeutung von CISM und anderer Faktoren für die eigene Erholung

Details zur Zielsetzung und Methodik der einzelnen Beschreibungsmerkmale:

Es wurde weitestgehend mit einer Prozentskala gearbeitet. Der Vorteil der Prozentskala wurde in der leichten Verständlichkeit, der einfachen Berechnung von Kennwerten und der vielseitigen Anwendung gesehen. Je nach Merkmalsbereich konnte die Skala mit unterschiedlichen Ausprägungsmerkmalen (Intensität, Zustimmung, Häufigkeit, ...) versehen werden.

1. Neben einer offenen Beschreibung des kritischen Ereignisses und der unmittelbaren Reaktionen wurde auch nach dem Ausmaß an Beunruhigung/Aufregung gefragt (Prozentskala).
2. Die Reaktion auf das kritische Ereignis durch das Team und die Führung wurde hier (in Ergänzung zu (7)) zunächst offen erfragt, um möglichst die gesamte Bandbreite an Interventionen und Unterstützungsmaßnahmen zu erfahren.

3. Die Frage nach der Dauer der Leistungseinschränkung war wichtig, um die in (5) und (7) erfragten Einschätzungen zu quantifizieren und den *Return on Investment* des CISM-Programms berechnen zu können.
4. Für die Darstellung der *Performance*-Kurve wurden gegenüber der Machbarkeitsstudie (Kapitel 2.2.2) zwei wesentliche Änderungen vorgenommen. Einerseits wurden feste Zeitintervalle in Wochen vorgegeben: Die durchschnittliche Leistung wurde für die Woche vor dem kritischen Ereignis (Basislinie), für die Woche des kritischen Ereignisses und die sechs folgenden Wochen angegeben. Hier sollten entsprechend Wochendurchschnittswerte eingetragen werden. Andererseits wurde eine Differenzierung der Leistung in die Arbeitslast und die Bewältigung dieser Arbeitslast eingeführt. Die Arbeitslast ist für Fluglotsen ein gebräuchliches Konstrukt. Sie ist im Kontext der Lotsenarbeit wichtig, da sie quasi den „Arbeitstakt“ darstellt. Niedrige Leistung kann sowohl das Ergebnis gut bewältigter geringer Arbeitslast sein, als auch das Ergebnis schlecht bewältigter großer Arbeitslast. Die Differenzierung bietet die Möglichkeit, die Bewältigung der Arbeitslast als eigenständigen Indikator für die Leistung zu untersuchen und die Arbeitslast als Situationsvariable einzubeziehen.
5. Um die Veränderung in Bezug auf Fähigkeiten, Tätigkeitsausübung und Leistungsergebnisse zu erfassen, wurden die Befragten gebeten, die Ausprägung für den Belastungszeitraum in die gleiche Skala einzutragen, in die sie bereits ihre Basislinie eingetragen hatten. So wurde gewährleistet, dass die Fluglotsen die eigene Basislinie als Referenz nutzten und Verbesserungen oder Verschlechterungen bewusst angaben.
6. Die Erfassung der emotionalen und körperlichen Beeinträchtigungen erfolgte anhand von zwei etablierten Instrumenten: Dem Gießener Beschwerdebogen (Brähler & Scheer, 1983) sowie Trainingsmaterialien der *International Critical Incident Stress Foundation* ICISF, die eine Symptomcheckliste enthalten. Die aus beiden Instrumenten zusammengestellten Beschwerden wurden hinsichtlich ihrer Intensität im Belastungszeitraum auf einer 5er-Skala eingeschätzt (0 nicht, 1 kaum, 2 einigermaßen, 3 erheblich, 4 stark aufgetretene Beschwerden).
7. Hier sollten die Auswirkungen beschrieben werden, die durch die *Performance*-Kurve nicht erfasst werden, da diese eine abstrakte Leistung erfasst, ohne dass die impliziten Leistungsmerkmale klar sind. Daher wurde hier differenziert nach den „*Key Perfor-*

mance Indicators“ für den Wertschöpfungsprozess „Flugsicherung“ gefragt: *Safety*, Kapazität und Flüssigkeit der Verkehrsabwicklung (bestimmt die Pünktlichkeit). Dies sind Merkmale für die Prozessebene im HPM (Kapitel 1.4), da sie die kooperative Leistung gemeinsam mit vor- und nachgelagerten Sektoren betreffen. Konkret wurde hierzu auch nach Folgen für das Zusammenwirken mit dem Planer (Fluglotse und Planer bearbeiten zusammen einen Sektor) und mit den Nachbarsektoren erfragt. Als weiteres Merkmal im Rahmen der Prozesseffektivität wurden operative Eingriffe durch die Führungskraft erfasst. Konkret wäre dies eine Veränderung der Einsatzplanung (z.B. Ausplanung oder Einsatzwechsel). Ergänzend wurde noch nach negativen Folgewirkungen auf die eigene Arbeit ohne Einfluss auf die Sicherheit gefragt. Der Fachterminus bei den Fluglotsen heißt „Arbeitsfehler“. Dazu zählt z.B. das Zurückfallen in die Muttersprache deutsch gegenüber deutschsprachigen Piloten, obwohl englisch vorgeschrieben ist; diese Abweichung ist nicht sachgemäß, obwohl sie in der Regel ohne Folgen bleibt.

8. Um den Einfluss von CISM zu konkretisieren wurde zunächst nach den individuellen Copingstrategien gefragt. Die Items orientierten sich einerseits an einem validierten Instrument (Stressverarbeitungsbogen von Janke, 1999) und andererseits an den in der CISM-Literatur beschriebenen Strategien zur Verarbeitung kritischer Ereignisse. Die Ausprägung wurde auch hier anhand der Prozentskala angegeben. Für die gleichen Items wurde nach der Unterstützungsleistung von CISM gefragt. So konnte analysiert werden, ob CISM die individuellen Bewältigungsstrategien unterstützt oder ergänzt. Um die Bedeutung dieser Unterstützungsleistung zu quantifizieren, wurde wie in der Machbarkeitsstudie (Kapitel 2.2.2) nach dem prozentualen Anteil von CISM und anderen Faktoren an der Erholung gefragt.

Abschnitt 3.: Erhebung der Einflüsse der Kommunikations- und Sicherheitskultur

Dieser Abschnitt diente der Erfassung der Kommunikationskultur, d.h. des Umgangs miteinander, und wie CISM diesen beeinflusst hat. Zunächst wurde die allgemeine Kommunikationskultur erfasst, dann die Kultur bezogen auf CISM und Sicherheit. Kommunikationskultur im Allgemeinen lässt sich beispielsweise beschreiben mit Adjektiven wie offen, flexibel, mitarbeiterorientiert, teamorientiert usw. Die Fluglotsen gaben an, inwieweit diese Adjektive auf die Kultur der DFS zutreffen (6-stufige Skala von „trifft genau zu“ bis „trifft gar nicht zu“). In einem zweiten Schritt wurde erfragt, ob CISM die Kultur bezogen auf jedes einzelne Adjektiv stark verbessert (++) , verbessert (+) , nicht verändert (0) , verschlechtert (-) oder stark verschlechtert (--) hat. Weitere mit der Prozentskala belegte Fragen erfassten, wie stark die Einführung von CISM das Management beeinflusst hat, inwieweit die Erfahrungen der Fluglotsen und *Peers* bei der Weiterentwicklung des Programms berücksichtigt wurden, wie nachhaltig die Führungskräfte CISM unterstützen und inwiefern das CISM-Programm der DFS als *Best Practice* für andere Luftfahrtorganisationen gesehen wird. Nachdem Kommunikationskultur und CISM bewertet wurden, gaben die Fluglotsen noch an, wie ausgeprägt sie die Sicherheitskultur auf verschiedenen Unternehmensebenen sehen (nicht, wenig, mittelmäßig, ziemlich, sehr ausgeprägt), welche Beiträge sie selbst und welche die Führung zur Sicherheitskultur leisten müssen. Zur Kontrolle wurde zuletzt gefragt, inwieweit die Innovationen der letzten Jahre (CISM und andere HF/HR/T-Programme) die Sicherheitskultur verbessert, unberührt gelassen oder verschlechtert haben und ob die Kulturfragen im Hinblick auf aktuelle Vorgänge in der Organisation (z.B. Zusammenlegung von Betriebsstätten) beantwortet wurden.

Abschnitt 4.: Erhebung der langfristigen Wirkungen von kritischen Ereignissen

Im vierten Abschnitt sollten die Fluglotsen ihr schwerwiegendstes kritisches Ereignis aus der Vergangenheit beschreiben. Es sollte länger als zwei Jahre zurückliegen, um einerseits eine Doppelung mit dem Fragebogenabschnitt 2 zu vermeiden und andererseits die Folgen dieses Ereignisses Jahre später zu erfassen. Daher wurde hier gezielt nach dem Auftreten von so genannten „*déjà-vu*-Situationen“ gefragt, d.h. ob Erinnerungen an das vergangene kritische Ereignis in späteren Arbeitssituationen auftraten. Es sollte sowohl die Häufigkeit von *déjà-vu*-Situationen als auch die damit verbundenen emotionalen und körperlichen Beeinträchtigungen anhand einer Beschwerdeliste (Brähler & Scheer, 1983; ICISF Material) eingeschätzt werden. Ergänzend wurden die Fluglotsen zu den Arbeitsauswirkungen dieser Beeinträchtigungen gefragt. Aussagen hierzu sollten Erkenntnisse darüber liefern, ob nicht mit CISM behandelte kritische Ereignisse bei Fluglotsen ggf. Jahre später noch zu Leistungseinschränkungen führen, was die Notwendigkeit der Unterstützung ggf. auch über den Er-

eigniszeitraum hinaus indiziert. Auch hier sollte der Anteil von CISM und anderer Faktoren an der Erholung eingeschätzt werden.

Abschnitt 5.: Erhebung der Wirkung eines aktuellen kritischen Ereignisses

Im fünften Abschnitt wurden die Fluglotsen gebeten, ein aktuelles kritisches Ereignis zu beschreiben. Dazu wurde der Abschnitt 5 als eigenständiges Dokument an die Fluglotsen gegeben, dass sie innerhalb der nächsten vier Monate ausfüllen sollten, falls sie von einem kritischen Ereignis betroffen sein würden. Obwohl klar war, dass hier nur eine relativ kleine Stichprobe zu erzielen war, wurde aufgrund der hohen Datenqualität eines „Echtzeit-Protokolls“ dieser zusätzliche Erhebungsaufwand investiert. Der Vergleich von aktuellen und zurückliegenden *Performance*-Kurven sollte darüber hinaus den Nachweis liefern, dass keine signifikanten Unterschiede im berichteten Verlauf bestehen und die retrospektive Einschätzung von Fluglotsen eine reliable und valide Erfassungsmethode ist. Die Korrelation der generierten *Performance*-Kurven von vergangenen und aktuellen Ereignissen war mit 0,91 ($p < 0,01$) in der Tat sehr hoch, so dass man davon ausgehen kann, dass auch die Beschreibung der Leistungsfähigkeit aus dem Gedächtnis heraus (Abschnitt 2) eine hohe Aussagekraft hat.

Der Fragebogen des Abschnittes 5 war identisch mit dem Bogen des Abschnittes 2. Differenzierter wurde allerdings die *Performance*-Kurve erhoben, weil hier zusätzlich eine tagesgenaue Angabe der Arbeitslast und der Bewältigung der Arbeitslast über die ersten 14 Tage gefordert wurde. Der folgende Zeitraum 3-6 Wochen nach dem Ereignis wurde wiederum wochenweise erfasst.

2.2.5.4 Fragebogen für Peers

Der Fragebogen für die *Peers* zielte zum einen darauf ab, die Auswirkungen eines kritischen Ereignisses und die Wirkung von CISM auch aus der Fremdperspektive zu betrachten, um Hinweise zur Validität der Selbsteinschätzung zu gewinnen. Zum anderen ging es um die Untersuchung der Wirksamkeitsfaktoren von CISM, die in der Kompetenz der *Peers* und dem CISM-Prozess begründet sind. Entsprechend wurde der *Peer*-Fragebogen z.T. analog zum Fluglotsen-Fragebogen konzipiert.

Abschnitt 1.: Erhebung der *Peer*-Fähigkeiten und Ergebnisse der *Peer*-Tätigkeit

Dieser Abschnitt sollte vor allem Hinweise für die *Peer*-Ausbildung liefern und damit die Arbeit des CISM-Verantwortlichen unterstützen. Die *Peers* konnten sich anhand der 12 zentralen Fähigkeiten eines kollegialen Beraters, der 10 wichtigsten Tätigkeitsbereiche und anhand von 10 Ergebnisfaktoren selbst einschätzen. Diese Daten waren für die Wirtschaftlichkeitsanalyse nicht relevant.

Abschnitt 2.: Beschreibung einer Betreuung aus den letzten zwei Jahren

Vergleichbar zum Lotsenfragebogen sollten die *Peers* einen Betreuungsfall aus den letzten zwei Jahren einschätzen. Analog dazu wurden die Beeinträchtigungen des Fluglotsen, dessen Leistungsfähigkeit, die erlebten Auswirkungen auf seine Arbeit, die Bewältigungsmechanismen sowie Form und Bedeutung der Unterstützung durch CISM erfragt. Um zur *Performance*-Kurve des Fluglotsen annähernd vergleichbare Daten aus der Fremdperspektive zu generieren, wurden die *Peers* um ihre Einschätzung der Leistungsfähigkeit des Fluglotsen zu den Zeitpunkten der CISM-Gespräche gebeten. Aus den punktuellen Daten sollte eine *Performance*-Verlaufskurve errechnet werden. Die Korrelation dieser Fremdeinschätzung der *Peers* mit der Selbsteinschätzung der Fluglotsen sollte letztere als gültige Erfassungsmethode zusätzlich absichern. Darüber hinaus sollten die *Peers* anhand der in Abschnitt 1 bereits benutzten Fähigkeits- und Tätigkeitsmerkmale ihre Unterstützungsarbeit charakterisieren. Gedacht war dies als Validierung der allgemeinen Selbsteinschätzung. Die beiden vergleichbaren Kurvenabschnitte der Fremd- und Selbsteinschätzung korrelierten mit 0,89 sehr hoch ($p < 0,05$) miteinander und bestätigten sich so gegenseitig in ihrer Gültigkeit (*concurrent validity*).

Abschnitt 3.: Erhebung der Einflüsse der Kommunikations- und Sicherheitskultur

Der Abschnitt 3 wurde vollständig aus dem Lotsenfragebogen übernommen.

Abschnitt 4.: Erhebung der langfristigen Wirkungen von kritischen Ereignissen

Im vierten Abschnitt sollten auch die *Peers* die Häufigkeit und Intensität von Nachwirkungen („*déjà-vu*-Situationen“) kritischer Ereignisse sowie deren Auswirkungen beschreiben, falls sie solche bei sich oder Kollegen beobachten mussten.

Abschnitt 5.: Beschreibung einer aktuellen Betreuung zwischen Juni – September 2005

Auch die *Peers* wurden gebeten ggf. ein aktuelles kritisches Ereignis und den Erholungsverlauf zu protokollieren, falls sie im Erhebungszeitraum als kollegialer Berater tätig würden. Die Einschätzungsmerkmale waren analog zum Abschnitt 2 des Peerfragebogens, lediglich bei der *Performance*-Kurve wurde hier soweit möglich um eine tagesgenaue Einschätzung der Leistungsfähigkeit gebeten. Zusätzlich zur Leistungsfähigkeit sollte auch der Verlauf der emotionalen Beeinträchtigung aus der Sicht des *Peers* protokolliert werden. Der *Peer* wurde gebeten, mit dem betroffenen Fluglotsen abzusprechen, dass er seine Eindrücke dokumentiert und den Fragebogen-Code des Fluglotsen auf seinem Fragebogen vermerkt, um einen direkten Vergleich von Selbst-

und Fremdeinschätzung zu ermöglichen ohne die Anonymität gegenüber den Auswertern der Fragebogen aufzuheben.

2.2.6 Stichprobenbeschreibung und Rücklauf

Die Befragung der Fluglotsen und *Peers* fand in der Zeit von Juni bis Oktober 2005 statt. Fünf *Radarkontrollzentrale* (Berlin, Bremen, Karlsruhe, Langen und München) sowie sechs *Towemiederlassungen* (Düsseldorf, Hamburg, Hannover, München, Nürnberg und Stuttgart) der Deutschen Flugsicherung beteiligten sich. Frankfurt *Tower* wurde nicht einbezogen, stattdessen aber der Bereich Frankfurt *Approach*. Um eine möglichst hohe Beteiligung der Fluglotsen zu bewirken, wurden *Peers* in den Niederlassungen gebeten, die Fragebogen persönlich an die Fluglotsen zu verteilen. Insgesamt wurden über diesen Weg 1.030 Fragebogen an Fluglotsen gegeben. Der Fragebogenrücklauf betrug 30%, d.h. 309 Fragebogen wurden von den Fluglotsen ausgefüllt und zurückgeschickt. Beschrieben wurden von 77 Fluglotsen vergangene, kritische Ereignisse (Abschnitt 2) und von 10 Fluglotsen aktuelle, kritische Ereignisse im Erhebungszeitraum (Abschnitt 5). Für die statistischen Auswertungen wurden die aktuellen und die zurückliegenden Ereignisse in einem Datensatz vereinigt, die tageweise erfassten *Performance*-Kurven der aktuellen Berichte wurden zu diesem Zweck zu Wochenwerten zusammengefasst. Die genauere Aufschlüsselung des Rücklaufes nach Niederlassungen zeigt Abbildung 7.

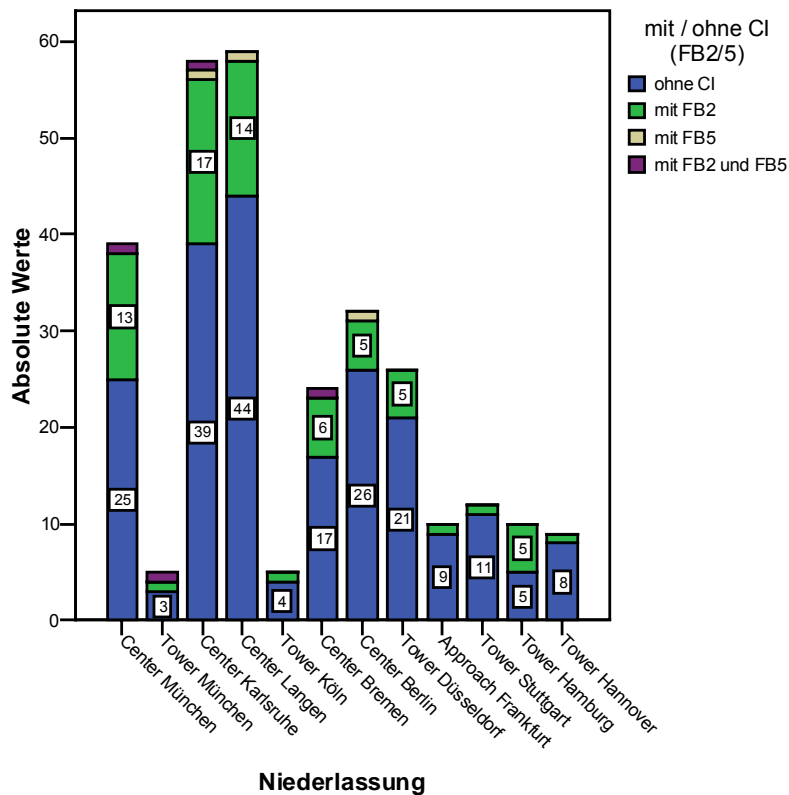


Abbildung 7: Fragebogenrücklauf nach Niederlassungen (CI: *Critical Incident*, FB: Fragebogenabschnitt)

51 von diesen 87 Fluglotsen mit zurückliegendem oder aktuellem Ereignis haben ein CISM-Gespräch in Anspruch genommen. Eine genauere Analyse der kritischen Ereignisse anhand der offenen Antworten zeigte allerdings, dass 21 Beschreibungen nicht die typischen Merkmale eines *Critical Incident* (CI) und den damit verbundenen *Critical Incident Stress* Reaktionen aufweisen (Leonhardt & Vogt, 2006), so dass in die engere Auswertung der kritischen Ereignisse 66 Fragebogen eingingen. Ausgeschlossen wurden beispielsweise Ereignisse, die anderen Fluglotsen zugeschrieben wurden und bei denen lediglich leichter Ärger über die daraus resultierende Situation (Notwendigkeit der Anpassung der Verkehrsregelung) beschrieben wurde. Mit dieser verkleinerten Stichprobe konnte immer noch eine solide Basis für die Wirtschaftlichkeitsbewertung von CISM geschaffen werden, da ein hoher Anteil der realen kritischen Ereignisse bei der DFS im untersuchten Zeitraum damit erfasst wurde.

Bei der Befragung der *Peers* erhielten alle *Peers* in den einbezogenen 11 Niederlassungen je einen Bogen. Insgesamt wurden 43 Bogen ausgegeben. Der Rücklauf betrug hier 100% (N = 43). Darin wurden 35 kritische Ereignisse aus den vergangenen zwei Jahren und sieben kritische Ereignisse aus dem Untersuchungszeitraum beschrieben. Eine detaillierte Aufschlüsselung des Rücklaufes nach Niederlassungen zeigt Abbildung 8.

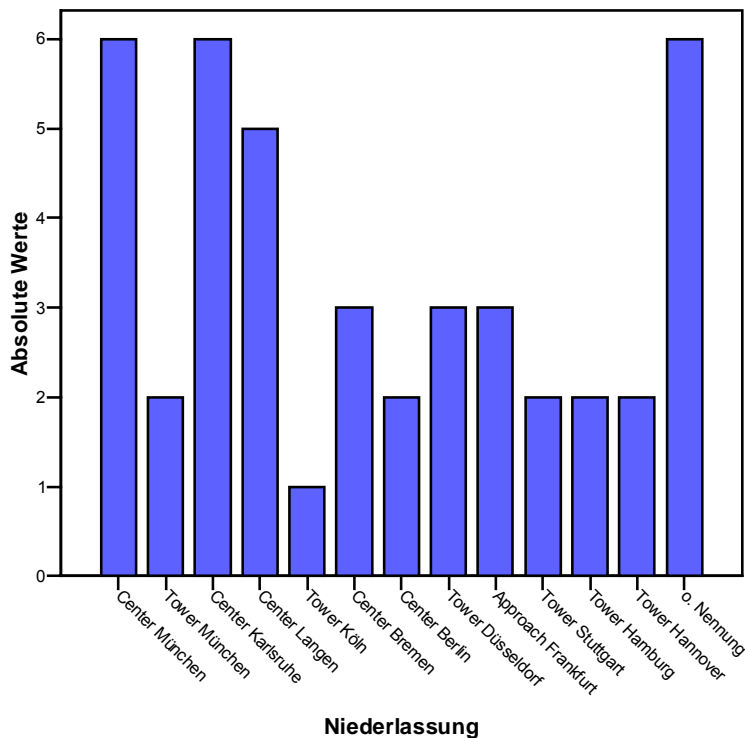


Abbildung 8: Erhaltene *Peer*-Fragebogen nach Niederlassungen

2.2.7 Auswertung der Fragebogen

Die Auswertung der Fragebogen erfolgte gemäß dem Modell der drei Steuerungskreisläufe (Abbildung 6) in folgenden Schritten:

1. Zunächst wurden die **Selbststeuerung des** Fluglotsen nach einem kritischen Ereignis und die direkten Auswirkungen des Ereignisses auf dessen Leistungsfähigkeit untersucht. Darüber hinaus wurden die Fluglotsen gebeten, die spezifischen Einflüsse von CISM zu berichten, die sie erlebt haben. Zur Überprüfung konnten hier die Angaben der *Peers* herangezogen werden.
2. Im zweiten Schritt wurde der **Einfluss der operativen Steuerung** geprüft: Wann und wie wurde von den Führungskräften eingegriffen und welche Konsequenzen hatte dies auf die Leistung und Erholung des betroffenen Fluglotsen?
3. Im dritten Schritt wurden die Wirkungen von kritischen Ereignissen und der getroffenen Steuerungsmaßnahmen auf die strategischen Zielgrößen und die **Bedeutung dieser Wirkungen für die strategische Steuerung auf Unternehmensebene** betrachtet.
4. Die Auswertung von Abschnitt 3 (Sicherheitskultur) und Abschnitt 4 (langfristige Folgen) der Fragebogen erfolgte zunächst unabhängig von den Steuerungskreisläufen. Es galt, langfristige, nicht-monetäre und systemische Wirkungen von CISM aufzuzeigen, die das Bild der Wirtschaftlichkeitsbetrachtung abrunden.

2.2.7.1 Selbststeuerung der Fluglotsen

Die Selbststeuerung wurde im Hinblick auf folgende Fragen analysiert:

1. Wie stark war die psychische Beeinträchtigung durch das kritische Ereignis?
2. Wie lange hielt die Beeinträchtigung an?
3. Welcher Art (emotional, körperlich, verhaltensbezogen) war die Beeinträchtigung?
4. Wie sehr war dadurch die generelle Leistungsfähigkeit betroffen?
5. Wie haben die Fluglotsen versucht, das kritische Ereignis zu bewältigen und ihre Arbeitsfähigkeit wieder herzustellen bzw. zu stabilisieren?
6. Wie viele Fluglotsen nahmen CISM in Anspruch?
7. In welcher Form nutzten die Fluglotsen CISM zur Unterstützung?
8. Wie wirkte sich das Ereignis auf die einzelnen Fähigkeiten und Tätigkeiten aus und wurden arbeitsbezogene Beeinträchtigungen durch CISM vermieden?
9. Konnte CISM die Arbeitsqualität und Produktivität nach einem kritischen Ereignis stabilisieren?

Wie stark war die psychische Beeinträchtigung durch das kritische Ereignis?

Zunächst wurden die Fluglotsen gefragt, wie sehr das kritische Ereignis sie beunruhigt bzw. aufgeregt hat (auf einer Skala von 0 – 100). Dieses Globalitem geht auf den *Emotional Response to Stress Test* (Turner & Naismith, 1989) zurück, dessen Item *upset* in Absprache mit den Autoren explizit mit beiden deutschen Bedeutungen beunruhigt/aufgeregt belegt ist. Es gilt als Maß für die empfundene emotionale Betroffenheit, die das Ereignis ausgelöst hat. Abbildung 9 zeigt die Verteilung des Ausmaßes an Beunruhigung.

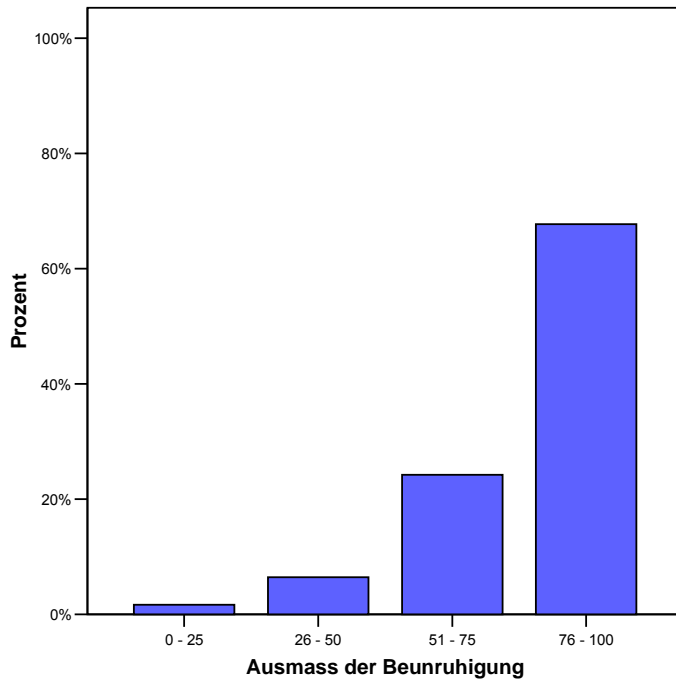


Abbildung 9: Prozentanteil der Fluglotsen mit geringer (0-25% und 26-50%), moderater (51-75%) und starker (76-100%) Beunruhigung/Aufregung durch ihre kritischen Ereignisse

Fast alle Fluglotsen nannten ein Ausmaß an Beunruhigung/Aufregung von über 50%. Der Durchschnittswert lag bei 82,3%. Dies zeigt, dass die (verkleinerte) Stichprobe Ereignisse umfasst, die von den betroffenen Fluglotsen in hohem Maße als kritisch und emotional belastend erlebt wurden. Dies ist die zentrale Indikation für CISM (Leonhardt & Vogt, 2006). Die *Peers* beschrieben die von ihnen erlebte Beunruhigung/Aufregung ähnlich stark mit einem Mittelwert von 75%. Hier lagen 80% der Angaben im Bereich von über 50%iger Beunruhigung/Aufregung.

Wie lange hielt die Beeinträchtigung an?

Der angegebene Zeitraum von erlebter Beeinträchtigung betrug durchschnittlich 10,8 Tage (Belastungszeitraum). Die Abbildung 10 zeigt, dass fast ein Drittel der Fluglotsen angaben, dass sie 2-3 Tage belastet waren und ca. ein Viertel ihre Beeinträchtigung zwischen 4 und 7 Tagen ansiedelten.

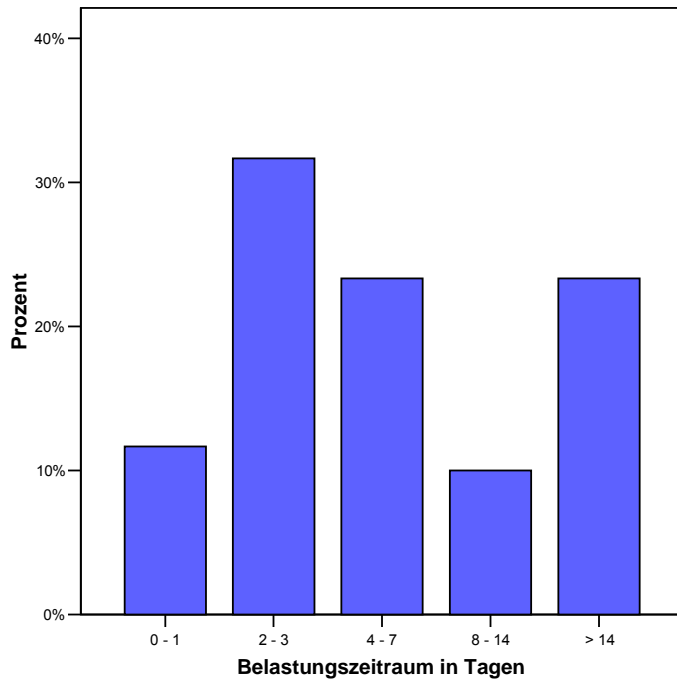


Abbildung 10: Von Fluglotsen angegebener Belastungszeitraum nach dem kritischen Ereignis

Die *Peer*-Einschätzung weicht davon leicht ab (s. Abbildung 11). Sie sahen mehr Fluglotsen (ca. 30%), die maximal einen Tag beeinträchtigt waren. Dies mag allerdings daran liegen, dass die Mehrheit der *Peers* den betroffenen Fluglotsen nur am ersten Tag des kritischen Ereignisses erlebten; nur 16 *Peers* (38%) haben den Fluglotsen zu einem späteren Zeitpunkt betreut und konnten daher einen fundierten Eindruck der Belastung über den ersten Tag des kritischen Ereignisses hinaus geben.

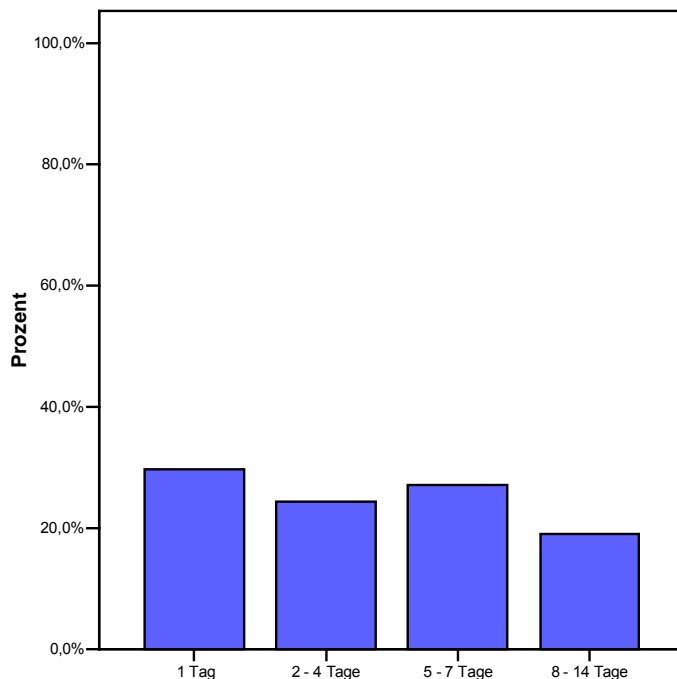


Abbildung 11: *Peer*-Einschätzung der Belastungsdauer nach dem kritischen Ereignis

Welcher Art war die verhaltensbezogene, emotionale und körperliche Beeinträchtigung?

In der Befragung wurden Teile des Gießener Beschwerdebogens, Teile einer ICISF-Symptomcheckliste und selbstkonstruierte Items eingesetzt, um die Art der verhaltensbezogenen, emotionalen und körperlichen Beeinträchtigung zu erfassen. Die Ergebnisse zeigt Abbildung 12. Die Einschätzung bezieht sich (wie bei allen folgenden Werten) auf den gesamten Belastungszeitraum.

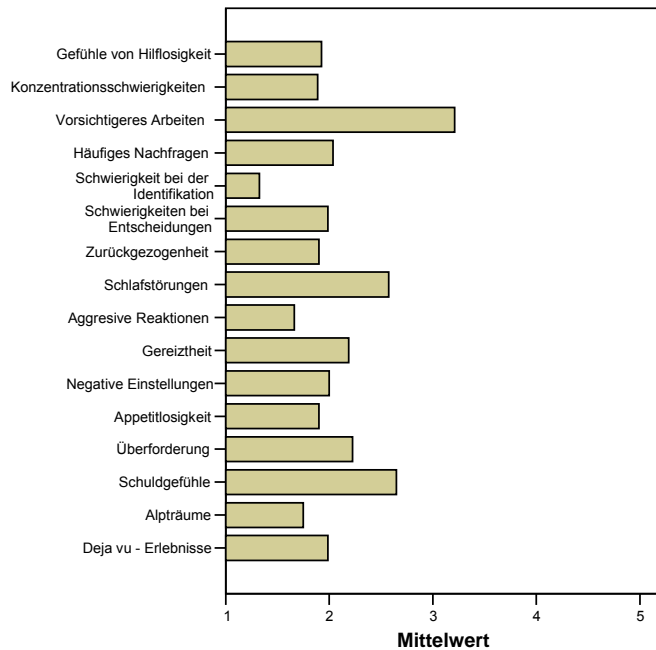


Abbildung 12: Art und Ausmaß der verhaltensbezogenen, emotionalen und körperlichen Beeinträchtigung (0 nicht, 1 kaum, 2 einigermaßen, 3 erheblich, 4 stark aufgetretene Beschwerden)

Die deutlichsten Reaktionen der Fluglotsen sind Ängstlichkeit („vorsichtigeres Arbeiten“) verbunden mit Schuldgefühlen, Schlafstörungen und dem Gefühl der Überforderung. Diese drei Symptome stehen für eine deutliche Verunsicherung des Fluglotsen nach dem kritischen Ereignis. Aus dieser Befindlichkeit heraus trauten sie sich die Bewältigung der normalen Arbeitslast nach dem kritischen Ereignis nicht mehr ohne weiteres zu und drosselten ihr Arbeitstempo, um die Sicherheit weiter zu gewährleisten. Dies wurde von vielen Fluglotsen als direkte Reaktion in den offenen Antworten beschrieben. Schlafstörungen signalisieren, dass die Fluglotsen das Ereignis immer wieder beschäftigt und „nicht in Ruhe lässt“. Die emotionale und kognitive Verarbeitung des Vorfalls findet permanent statt und benötigt viele eigene Ressourcen. Die Peers schätzten die Beeinträchtigungen der Fluglotsen im Profil ähnlich ein und bestätigten damit die Lotsenwerte (s. Abbildung 13). Die Ausprägung wurde sogar ten-

denziell, insbesondere bei den Schuldgefühlen und der erlebten Überforderung, noch höher eingeschätzt.

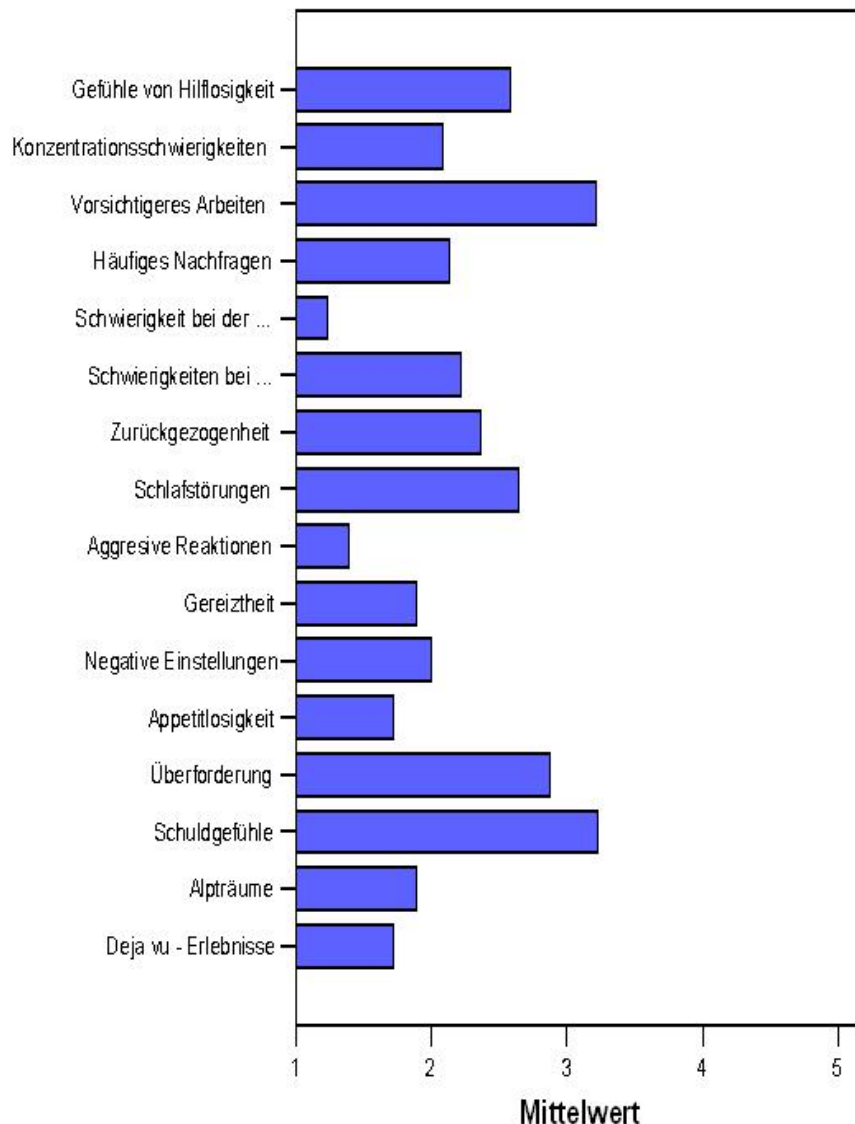


Abbildung 13: Von den *Peers* berichtete Beeinträchtigungen der Fluglotsen (0 nicht, 1 kaum, 2 einigermaßen, 3 erheblich, 4 stark aufgetreten)

Möglicherweise erlebten die *Peers* in ihrer Rolle als „Helfer“ die Hilflosigkeit bei den betroffenen Fluglotsen besonders stark.

Wie sehr war dadurch die generelle Leistungsfähigkeit betroffen?

In der Woche nach dem kritischen Ereignis war die erlebte generelle Leistungsfähigkeit um ca. 20 – 25% reduziert. Dies zeigt die von den Fluglotsen für die Zeit vor und nach dem kritischen Ereignis angegebene *Performance*-Kurve (= Bewältigung der wahrgenommenen Arbeitslast). Dabei wurden die z.T. tagesgenauen Werte für die aktuellen kritischen Ereignisse aus Gründen der Vergleichbarkeit zu Wochenwerten aggregiert. Die eingeschränkte Leistungsfähigkeit führte nur in zwei Fällen zu realen Ausfallzeiten (*day-off*). In allen anderen Fällen gingen die Fluglotsen in den Tagen nach dem kritischen Ereignis wieder ihrer Lotsentätigkeit nach. Die „*day-offs*“ wurden

aus der Berechnung der Leistungskurve herausgenommen, da nicht mit Sicherheit geschlossen werden kann, in welchem Umfang an diesen Tagen eine Leistungsmin- derung vorlag.

Resultiert aus einer Beeinträchtigung statt realer Ausfallzeiten eine Leistungsminde- rung am Arbeitsplatz, so spricht man von Präsentismus („*ill at work*“). Der Begriff wurde geprägt in Ergänzung zum Begriff des Absentismus. Präsentismus trägt der Beobachtung Rechnung, dass selbst anwesende Mitarbeiter häufig nicht ihre volle Leistung zeigen. Karch (2001) berichtet in einem Übersichtsartikel, dass Mitarbeiter im Durchschnitt nur 65-70% ihres Leistungspotentials zeigen. Die Gründe hierfür können in der Person des Mitarbeiters liegen, der z.B. krank zur Arbeit kommt, durch private oder berufliche Sorgen abgelenkt ist. Häufig ist Präsentismus aber auch ein Zusammenspiel aus individuellen, situativen (z.B. mangelhafte technische Unterstüt- zung) und organisationalen Faktoren (z.B. Führung, Organisationskultur, leistungs- ungerechte Bezahlung).

Der Präsentismuseffekt in der CISM-Evaluation lässt sich in eine Ausfallzeit umrech- nen, wenn man in der Leistungskurve die Fläche unterhalb des für die Woche vor dem kritischen Ereignis erhobenen Normalwertes der Leistung (ca. 80%, rote Linie in Abbildung 14) berechnet. Der Wert betrug in den erfassten sieben Wochen im Durchschnitt 5,4 Tage pro Fluglotse.

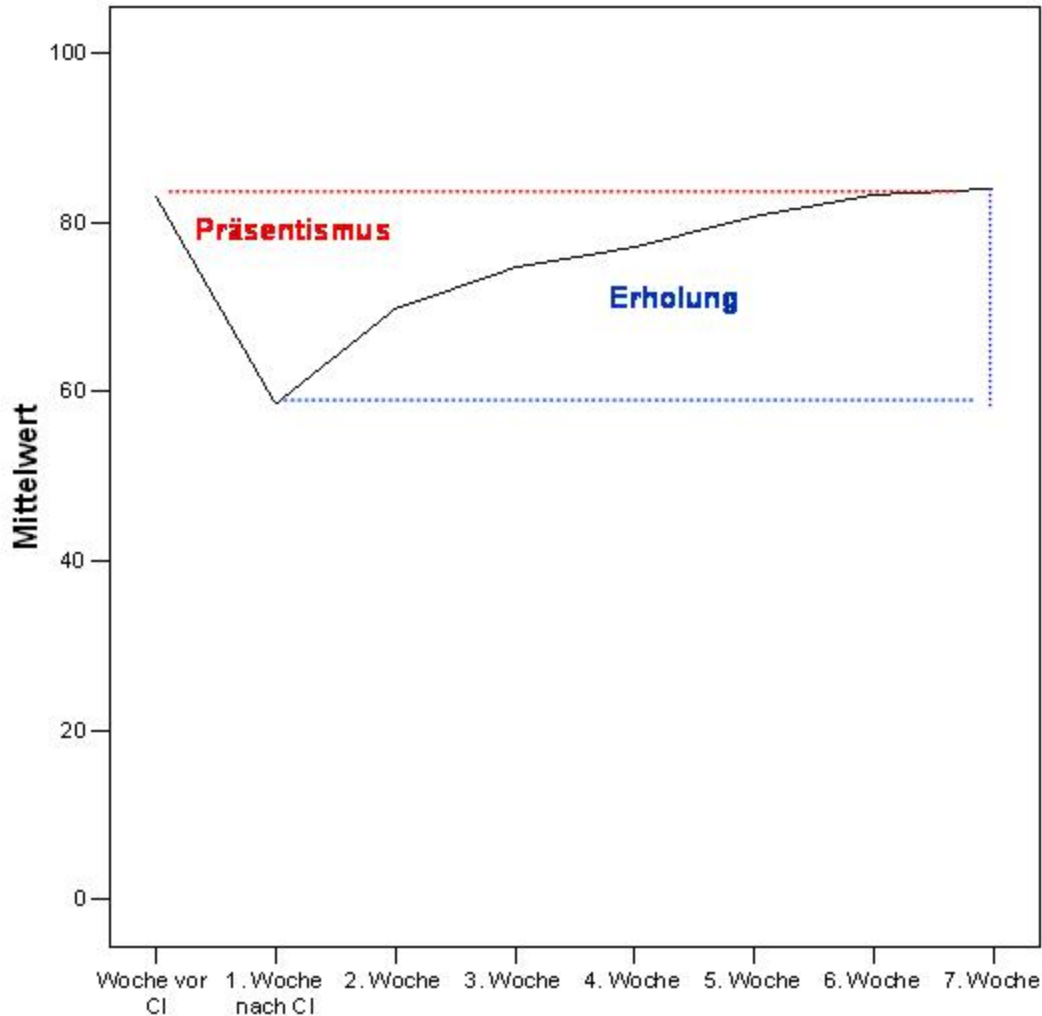


Abbildung 14: *Performance-Kurve* (Bewältigung Arbeitslast/Arbeitslast) der Fluglotsen vor und nach dem kritischen Ereignis (CI; *Critical Incident*)

Anhand der Leistungskurve ließ sich auch eine Erholungsfläche darstellen, die zwischen der blauen Linie und der *Performance-Kurve* in Abbildung 14 liegt. Die Erholung entspricht dann der Leistungsverbesserung in Bezug auf den minimalen Leistungswert nach dem kritischen Ereignis. Die Summe der zurück gewonnenen Leistungsfähigkeit innerhalb der dokumentierten sieben Wochen betrug 8,5 Tage. Diese gewonnenen Leistungstage konnten zum Teil CISM zugeschrieben werden (s.u.). Die Erholung fand vor allem in den ersten beiden Wochen nach dem Ereignis statt (steiler Abschnitt der Kurve).

Um die Plausibilität der Selbsteinschätzung zu überprüfen, wurde auch in dem *Peer-Fragebogen* nach einer Leistungseinschätzung (Arbeitslast und Bewältigung) für den betreuten Fluglotsen nach dem kritischen Ereignis zum jeweiligen Zeitpunkt des CISM-Gesprächs gefragt. Bei mehreren Gesprächen wurden entsprechend mehrere Werte angegeben. Aus den punktuellen Daten konnte eine z.T. tagesgenaue Verlaufskurve für die ersten Tage nach dem kritischen Ereignis konstruiert werden (s. Abbildung 15). Allerdings fehlen Datenpunkte für die Zeit vor dem kritischen Ereignis und die Zeit 2-7 Wochen danach, da die *Peer-Konsultationen* i.d.R. nur am Tag des kritischen Ereignisses stattfinden.

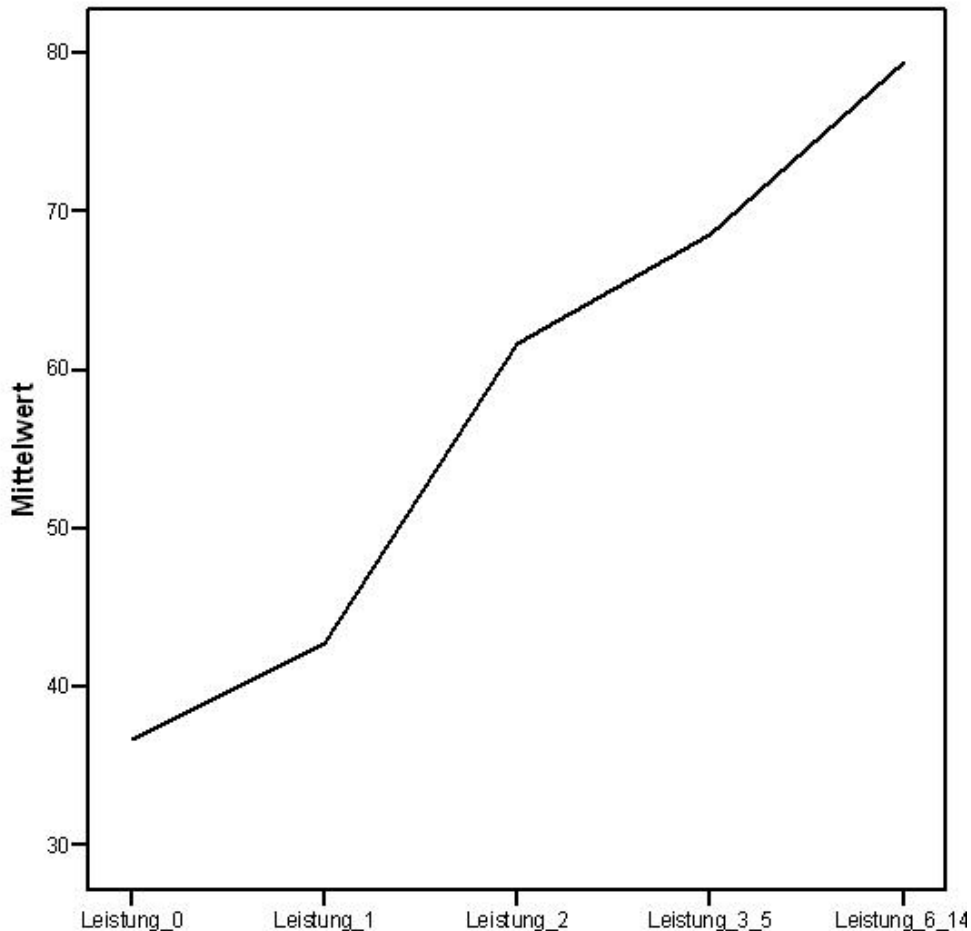


Abbildung 15: Leistung der durch *Critical Incidents* betroffenen Fluglotsen in Prozent (*Performance-Kurve*) aus Sicht der *Peers* in den ersten Tagen nach dem Ereignis
 (Leistung_0 = Tag des *Critical Incident*, Leistung_1= 1 Tag nach CI; Leistung_2= 2 Tage nach CI
 Leistung_3_5= 3-5 Tage nach CI; Leistung_6_14= 6-14 Tage nach CI)

Um genügend Daten für eine Mittelwertbildung zu integrieren, wurden die Zeiträume von 3 - 5 und von 6 - 14 Tagen nach dem Ereignis zusammengefasst. Hier war die Anzahl von CISM-Gesprächen verständlicherweise geringer und damit auch die Anzahl von Datenpunkten zur Leistungsfähigkeit des betroffenen Fluglotsen. Aus Sicht der *Peers* war die Minderung der Leistung nach dem Ereignis noch stärker als in der Selbsteinschätzung der Fluglotsen. Der Wert lag bei 36% Leistungsfähigkeit in der *Peer-* gegenüber 60% in der Lotseneinschätzung. Wie die Fluglotsen selbst, erlebten auch die *Peers* die stärkste Erholung zwischen dem ersten und dem zweiten Tag nach dem Ereignis. Nach zwei Wochen war aus Sicht der *Peers* die Leistungsfähigkeit von 80% wieder erreicht, die im Verständnis der Fluglotsen einer typischen, durchschnittlichen Leistungsfähigkeit entspricht. Die Fluglotsen selbst schätzten ihre Leistungsfähigkeit nach zwei Wochen etwas geringer ein (70%) und erreichten die 80% erst nach sieben Wochen. Die *Peer-Kurve* legt die Vermutung nahe, dass die Fluglotsen ihre Leistungsfähigkeit in den ersten Tagen nach dem kritischen Ereignis überschätzen und im weiteren Verlauf leicht unterschätzen. Die Berechnung der Präsentismus- und Erholungskurve ist entsprechend als konservativ einzustufen, d.h. die in der Hauptstudie gefundene Wirtschaftlichkeit des Programms dürfte noch höher sein, da die unterschätzte Leistungseinbuße größer ist als die überschätzte.

Wie haben die Fluglotsen versucht, das kritische Ereignis zu bewältigen und ihre Arbeitsfähigkeit wieder herzustellen bzw. zu stabilisieren?

Die Abbildung 16 zeigt, welche Vorgehensweisen die Fluglotsen angewandt haben, um das kritische Ereignis und die damit verbundene Beeinträchtigung zu bewältigen. Es zeigte sich, dass eine Reihe sehr verschiedener kognitiver und verhaltensmäßiger Strategien von den Fluglotsen genutzt wurden.

Die sachliche Analyse des Ereignisses war die von den Fluglotsen bevorzugte Entlastungsstrategie (z.B. „Gründe klären“, „den Vorgang gedanklich durchgehen“). An zweiter Stelle wurden konkrete Aktivitäten, die Halt und Struktur geben, genannt (z.B. „aktiv sein und Sport treiben“, „normalen Tagesablauf durchhalten“). An dritter Stelle wurde die emotionale Bearbeitung des Vorfalls beschrieben (etwa „sich aussprechen“, „du kannst damit fertig werden“, „sich nichts vorwerfen“). Demgegenüber sahen die Fluglotsen keine Option darin, sich der Situation zu entziehen. Möglicherweise würde es als „eigene Schwäche“ erlebt, wenn der Fluglotse sich z.B. für mehrere Tage vom Dienst befreien lassen würde. Die Bewältigung wurde also eher „*on-the-job*“ gesucht bzw. in Verbindung mit der normalen Tätigkeitsausübung.

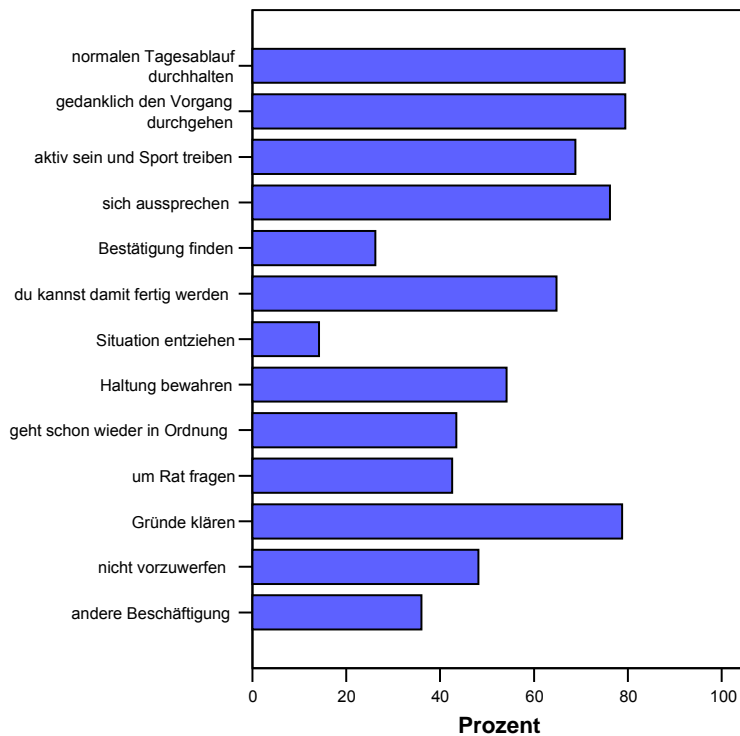


Abbildung 16: Bewältigungsmechanismen der Fluglotsen nach dem kritischen Ereignis

Diese Angaben wurden von den *Peers* bestätigt, die allerdings den kognitiven und kommunikativen Komponenten im Verhältnis zu den beiden Verhaltensstrategien etwas mehr Bedeutung gaben.

Wie viele Fluglotsen nahmen CISM in Anspruch und wie wirkte CISM?

Im Weiteren wurden die Daten hinsichtlich der Auswirkungen von CISM ausgewertet. In 48 der 66 Fälle mit kritischen Ereignissen wurde ein CISM-Gespräch in Anspruch genommen. Dies entspricht 73% aller Fälle. Ausgehend davon, dass CISM als Unterstützung von den Führungskräften der DFS empfohlen werden soll, kann man also feststellen, dass immerhin noch ein Viertel der Fluglotsen nicht erreicht wird bzw. die Führungskräfte ihrer diesbezüglichen Verantwortung nicht voll gerecht werden. Die Abbildung 17 schlüsselt die Nutzung von CISM nach Niederlassungen auf. Danach wurde CISM in einzelnen Niederlassungen stärker genutzt als in anderen. Dies ist vermutlich auf niederlassungsspezifische Faktoren zurückzuführen, wie z. B. unterschiedliche Kulturen oder Erfahrungen in Bezug auf CISM. Die Fluglotsen in Karlsruhe, die die Katastrophe von Überlingen miterleben mussten, zeigten das stärkste *Commitment* zum CISM-Programm. Diese Beobachtung wurde bereits in der Machbarkeitsstudie berichtet (Kapitel 2.2.2, siehe auch Vogt et al., 2004).

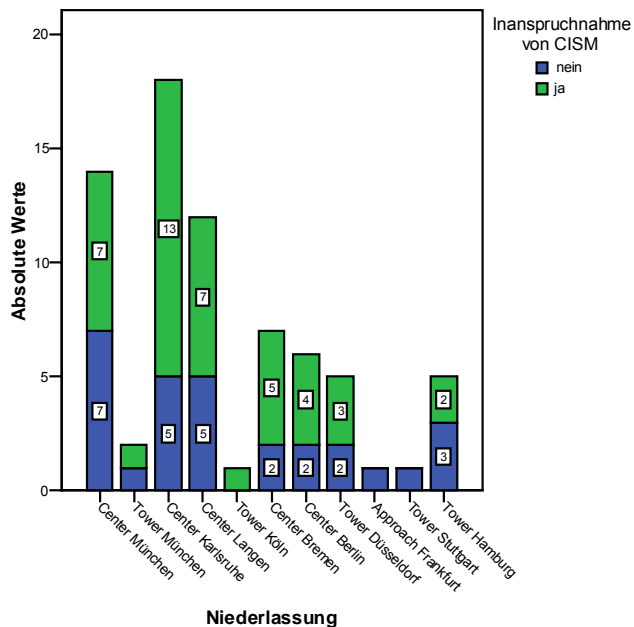


Abbildung 17: Nutzung von CISM in den verschiedenen Niederlassungen der DFS

In welcher Form nutzten die Fluglotsen CISM zur Unterstützung?

Um diese Frage zu klären, wurde zunächst einmal analysiert, ob die „CISM-Gruppe“ grundsätzlich auf andere Bewältigungsmechanismen zurückgriff, als die „Nicht-CISM-Gruppe“. Dazu wurde das bereits dargestellte Bewältigungsprofil getrennt für beide Gruppen ausgewertet. Es zeigte sich, dass es dort keine auffälligen Unterschiede gab, konkret waren keine Wertedifferenzen von über 10% feststellbar. D. h. dass die Fluglotsen generell auf die gleichen kognitiven, kommunikativen und verhaltensmäßigen Strategien zurückgriffen, egal ob sie eine Unterstützung von CISM in Anspruch nahmen oder nicht.

Darüber hinaus wurden die Fluglotsen der „CISM-Gruppe“ gefragt, wie stark CISM sie bei den verschiedenen Bewältigungsmechanismen unterstützt hat. Die Unterstützungsleistung von CISM in Bezug auf die generellen Bewältigungsmechanismen zeigt Abbildung 18.

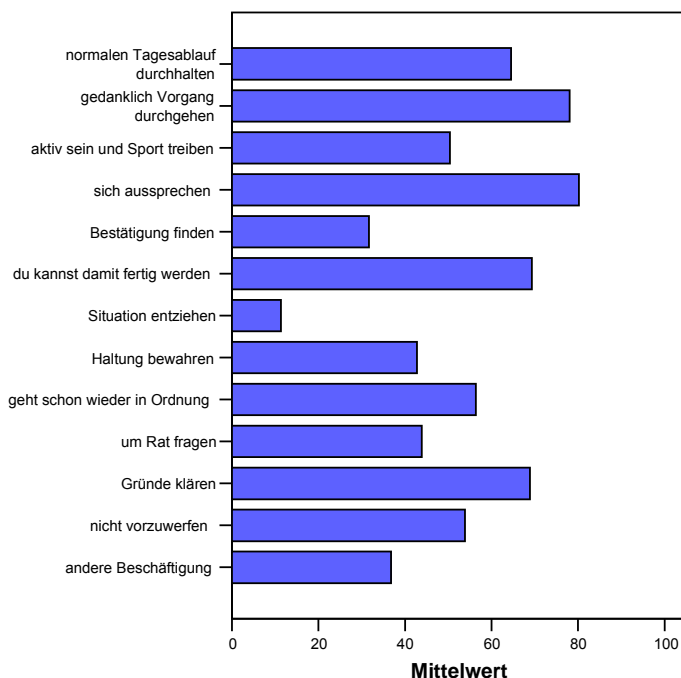


Abbildung 18: Ausmaß der Unterstützung der generellen Bewältigungsmechanismen durch CISM in Prozent

CISM unterstützte vor allem die eigenen, bevorzugten Bewältigungsmechanismen, welche wie bereits beschrieben in der sachlichen Klärung des Ereignisses und der Aussprache liegen. Die Peers förderten zusätzlich die präferierte Verhaltensstrategie, den normalen Tagesablauf durchzuhalten. Außerdem wurde eine gute Unterstützung durch CISM darin gesehen, Zuversicht in der Bewältigung des Ereignisses zu erhalten („Du kannst damit fertig werden“). Insgesamt zeigte sich, dass CISM keine neuen Strategien anbietet, die das eigene Spektrum ergänzen, sondern die bewährten „Hausmittel“ professionell verstärkt.

Die Peers sahen die eigene Unterstützungsleistung ebenfalls schwerpunktmäßig in diesen Bereichen (s. Abbildung 19).

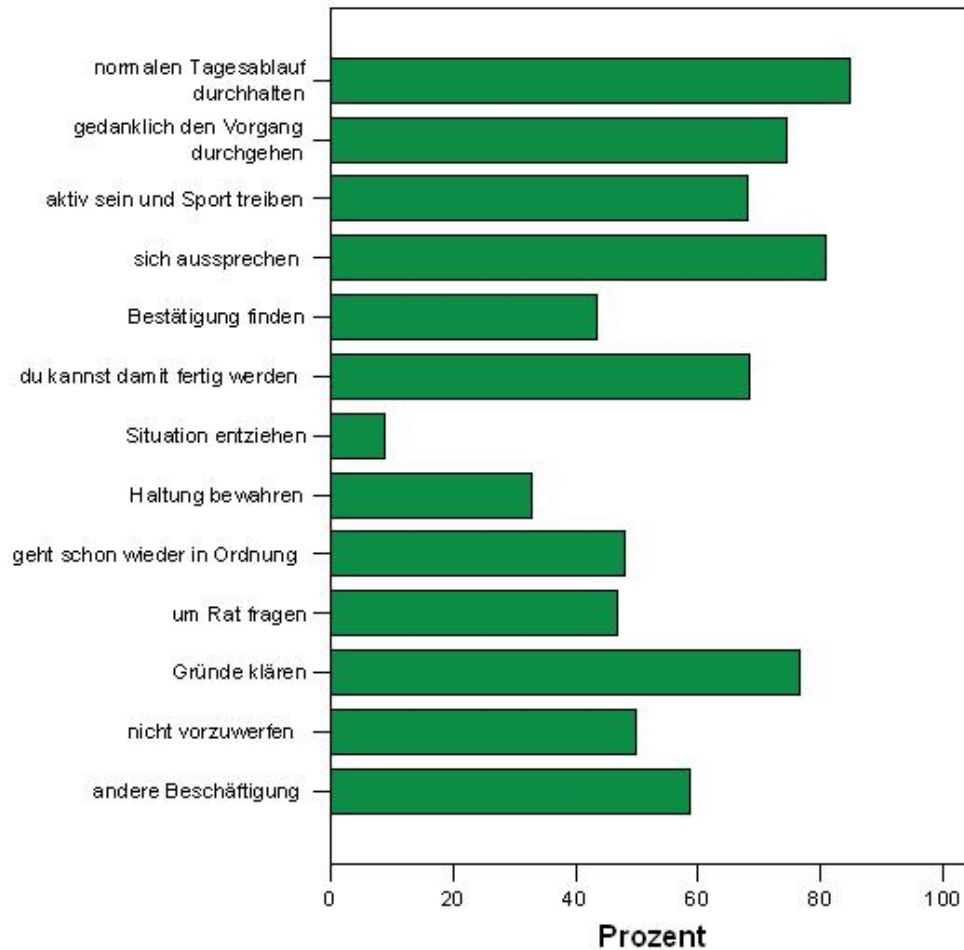


Abbildung 19: Unterstützung der generellen Bewältigungsmechanismen aus Sicht der *Peers* in Prozent

Für die *Peers* war allerdings eine die Fluglotsen unterstützende Verhaltensstruktur im Alltag noch wichtiger als für den Fluglotsen selbst. Dazu zählte neben dem normalen Tagesablauf und dem Sport auch die Suche nach anderen Beschäftigungen. Dieser letzte Punkt war von Fluglotsen als Unterstützungsleistung kaum genannt worden und spielte auch in der Einschätzung der eigenen Bewältigungsmechanismen eine geringe Rolle. Für die zukünftige Gestaltung der CISM-Interventionen ist zu prüfen, ob diese Unterstützungsform zielführend ist.

Neben der Art der Unterstützung durch CISM wurden die Fluglotsen auch wie bereits in der Machbarkeitstudie um eine Quantifizierung der Erholungswirkung von CISM und anderer Faktoren gebeten. Den prozentualen Beitrag, den CISM und andere Faktoren zur Erholung leisteten, zeigt Abbildung 20.

Das Ergebnis der Machbarkeitstudie wurden hinsichtlich der Bedeutung von CISM für die Erholung voll repliziert. Der Erholungsbeitrag des CISM-Gespräches wurde auf 36% geschätzt. In der Machbarkeitsstudie (Kapitel 2.2.2) waren es ebenfalls 36% (32%, wenn man die Katastrophe von Überlingen nicht berücksichtigt). Abweichend von den Ergebnissen aus der Machbarkeitsstudie wurde jedoch sowohl in der „CISM-Gruppe“ als auch der „Nicht-CISM-Gruppe“ die Bedeutung der Kommunikation mit Kollegen oder Freunden/Bekanntem mehr Gewicht gegeben als der Spontanerholung oder anderen Faktoren. Diese Einschätzungen wurden von den *Peers* geteilt, wobei der Erholungsanteil von CISM mit 42% in der Sicht der *Peers* noch etwas stärker zu Buche schlug.

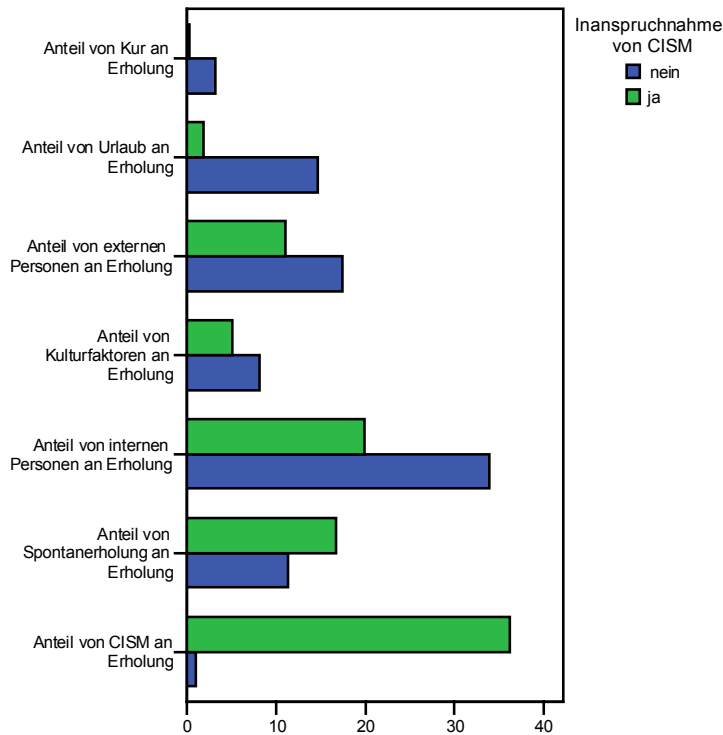


Abbildung 20: Beitrag der verschiedenen Einflussfaktoren zur Erholung in Prozent

Insgesamt setzte sich jedoch der Trend fort, der in der Machbarkeitsstudie (Kapitel 2.2.2) vermutet wurde: CISM fördert die Offenheit und Bereitschaft für Kommunikation über Emotionen und Verhalten auch bei den Fluglotsen, die noch nie ein CISM-Gespräch in Anspruch genommen haben. CISM hat in der DFS eine Kultur unterstützender Gespräche auch mit Kollegen geschaffen, die keine *Peers* sind.

Wie wirkt sich das Ereignis auf die einzelnen Fähigkeiten und Tätigkeiten aus und wurden job-relevante Beeinträchtigungen durch CISM vermieden?

Nachdem die bisherige Auswertung zeigte, dass kritische Ereignisse bei Fluglotsen mit z.T. erheblichen Beeinträchtigungen einhergingen, die sich in ihrer generellen Leistungsfähigkeit („*Performance-Kurve*“) niederschlugen, wurde nun der Frage nachgegangen, inwieweit dadurch auch spezifische Arbeitsfähigkeiten und Tätigkeitsausübungen betroffen waren. Um Veränderungen in den Fähigkeiten und der Tätigkeitsausübung nach dem kritischen Ereignis zu messen, wurden die Einschätzungen nach dem kritischen Ereignis mit der erhobenen Basislinie verglichen. Dazu wurden Differenzwerte zwischen beiden Angaben gebildet.

Die Veränderung der Fähigkeiten nach einem kritischen Ereignis im Vergleich der „CISM-Gruppe“ mit der „Nicht-CISM-Gruppe“ zeigt Abbildung 21.

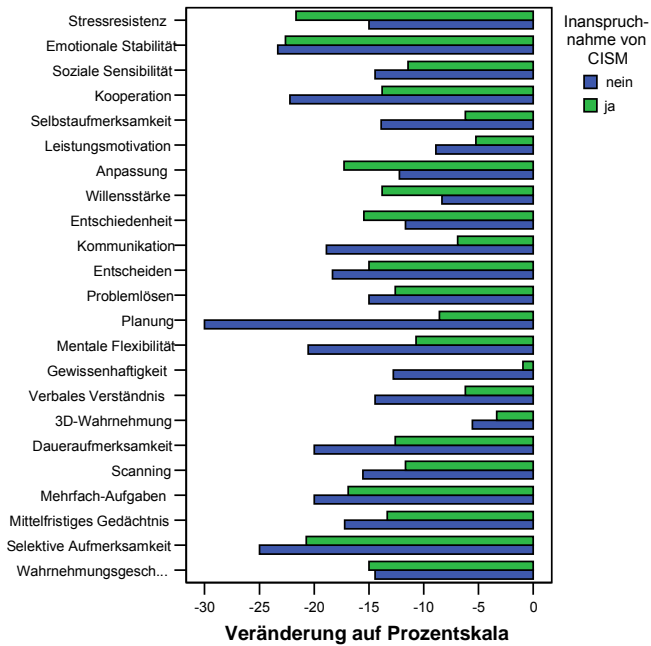


Abbildung 21: Veränderung der Fähigkeiten nach einem kritischen Ereignis in Prozent

Das Veränderungsprofil zeigt zunächst, dass alle Fähigkeiten in beiden Gruppen nach kritischen Ereignissen als vermindert eingeschätzt wurden. In beiden Gruppen ist insbesondere die emotionale Stabilität deutlich vermindert. Dies ist ein weiteres Indiz für die Verunsicherung der Fluglotsen nach dem Vorfall. Die Einschränkungen waren bei den meisten Fähigkeiten in der „Nicht-CISM-Gruppe“ größer. Besonders deutliche negative Veränderungen betrafen die für die Lotsenaufgabe zentralen kognitiven Fähigkeiten „Planung“, „Mentale Flexibilität“, „Selektive Aufmerksamkeit“ und „Daueraufmerksamkeit“. Diese waren bei den Fluglotsen, die keine CISM-Unterstützung erfuhren, deutlich stärker reduziert als bei den Fluglotsen, die ein CISM-Angebot wahrgenommen hatten. In ähnlicher Weise betroffen war die Kooperations- und die Kommunikationsfähigkeit der Fluglotsen ohne CISM.

Eine stärkere Auswirkung auf Fluglotsen, die CISM in Anspruch nahmen, fand sich hingegen bei vier sozialen Fähigkeiten. Zum einen waren dies die Stressstabilität und die Anpassungsfähigkeit. Dies drückt aus, dass sich die CISM-Fluglotsen vermindert in der Lage fühlten, nach einem kritischen Ereignis mit erneuten, hohen Anforderungen, z.B. einer neuen kritischen Flugverkehrssituation, umzugehen. Zum anderen waren die Fähigkeiten „Willensstärke“ und „Entschiedenheit“ betroffen, die beide bei den Fluglotsen ohne CISM nur wenig eingeschränkt waren. Es ergibt sich das Bild, dass der Fluglotse nach einem CISM-Gespräch deutlich selbstkritischer an seine weitere Arbeit geht und vor allem trotz seiner emotionalen Beeinträchtigungen nicht weniger gewissenhaft arbeitet. Er ist vorsichtiger in seinen Entscheidungen und der Einschätzung seiner eigenen weiteren Belastbarkeit. Der Fluglotse ohne CISM erkennt seine kognitiven Einschränkungen im Umgang mit dem komplexen System Flugsicherung, seine verminderte Leistungsfähigkeit im interaktiven System und seine Einbußen in der Gewissenhaftigkeit. Dennoch wertet er dies nicht in dem Maße als Einschränkung seiner Stressstabilität und Anpassungsfähigkeit wie der Fluglotse mit CISM.

Die Auswirkungen der verminderten Fähigkeiten konnten mit Hilfe der Tätigkeitsanalyse untersucht werden. Auch hier wurden Differenzwerte gebildet, um die Veränderung der Tätigkeitsausübung nach dem kritischen Ereignis zu messen. Die Veränderungswerte für alle untersuchten Tätigkeiten zeigt Abbildung 22.

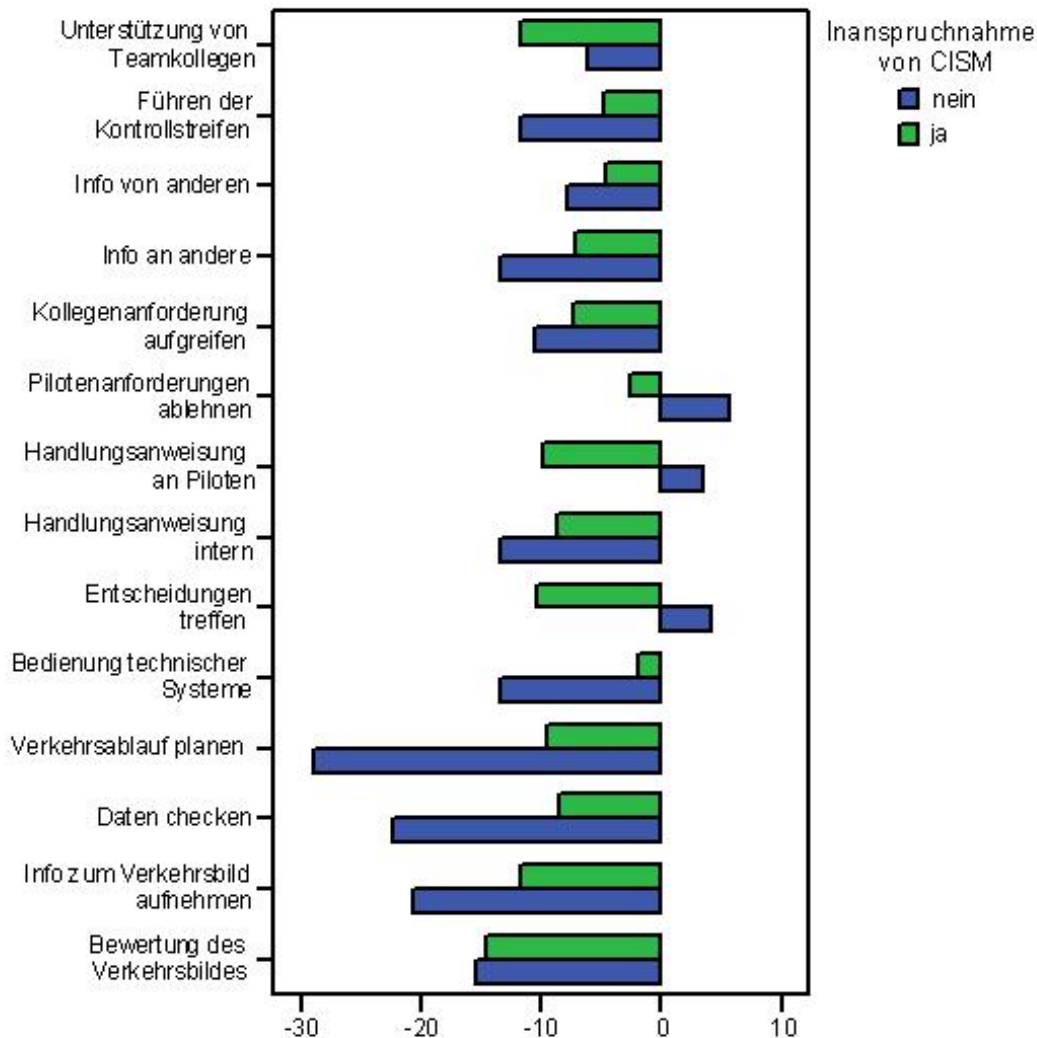


Abbildung 22: Veränderung der Tätigkeiten nach einem kritischen Ereignis in Prozent

Es zeigte sich hier eine deutliche Analogie zu den Fähigkeitseinschränkungen. Die eigene Tätigkeitsausübung wurde nach dem kritischen Ereignis in fast allen Bereichen als verschlechtert erlebt. In vielen Tätigkeiten erlebten sich die „Nicht-CISM-Fluglotsen“ als stärker eingeschränkt. Es waren vor allem die Tätigkeiten, die in engem Zusammenhang mit den eingeschränkten kognitiven Fähigkeiten zu sehen sind und die die Kerntätigkeit des Fluglotsen ausmachen: „Planung des Verkehrsablaufes“, „Daten checken“, „Informationen zum Verkehrsbild aufnehmen“ und „Bedienung technischer Systeme“.

Zum anderen waren kooperative und kommunikative Tätigkeiten bei den Fluglotsen ohne CISM eingeschränkt: „Information an andere geben“, „Kollegenanforderungen aufgreifen“ und „intern Handlungsanweisungen geben“. Gleichzeitig berichteten diese Fluglotsen, dass sie sich trotz ihrer Beeinträchtigung gut in der Lage fühlten, Entscheidungen zu treffen, die Piloten mit klaren Anweisungen zu versorgen und auch Anforderungen abzulehnen. Dies sind vor allem außenwirksame Tätigkeiten.

Fluglotsen, die ein CISM-Gespräch hatten, erlebten sich in diesen außenwirksamen Tätigkeiten dagegen eingeschränkt und waren zudem auch selbstkritischer, was ihre Möglichkeit anbetraf, Teamkollegen zu unterstützen.

Kann CISM die Arbeitsqualität und Produktivität nach einem kritischen Ereignis stabilisieren?

Die Veränderung im Leistungsprozess des Fluglotsen nach einem kritischen Ereignis wurde abschließend mit der Betrachtung der Leistungswirkungen analysiert. Es ging um die Frage, welche Auswirkungen diese verminderten Fähigkeiten und die Einschränkungen in der Tätigkeitsausübung auf die Arbeitsmenge und Arbeitsgüte haben. Dazu wurden die Fluglotsen befragt, zu welchen von sechs Ergebnisaspekten sich Auswirkungen ergeben haben. Die Einschätzungen sind in Abbildung 23 wiedergegeben.

Die deutlichsten Auswirkungen ergaben sich demnach auf die Verkehrsabwicklung, die Kapazität und die Zusammenarbeit. Die beiden letztgenannten Auswirkungen wurden vor allem von Fluglotsen genannt, die kein CISM-Gespräch genutzt hatten. Auswirkungen auf den für die Flugsicherung zentralen Ergebnisbereich Sicherheit wurden kaum genannt, insbesondere so gut wie gar nicht von den „CISM-Fluglotsen“. Bei den Fluglotsen ohne CISM-Gespräch nannten 11% solche Auswirkungen. Da *Safety* das herausragende Ziel der Flugsicherung ist, stellt die Minderung oder sogar Vermeidung von kleinsten Arbeitsfehlern, die die Sicherheit gefährden könnten, durch die CISM-Unterstützung einen wichtigen strategischen Nutzeneffekt dar.

Auch hier zeigt sich also, dass CISM Einschränkungen in der Leistung des Fluglotsen nach einem kritischen Ereignis deutlich vermindert. Dies betrifft sowohl die Arbeitsqualität (Verkehrsabwicklung; eingeschränkt auch die Sicherheit) als auch die Arbeitsmenge (Kapazität).

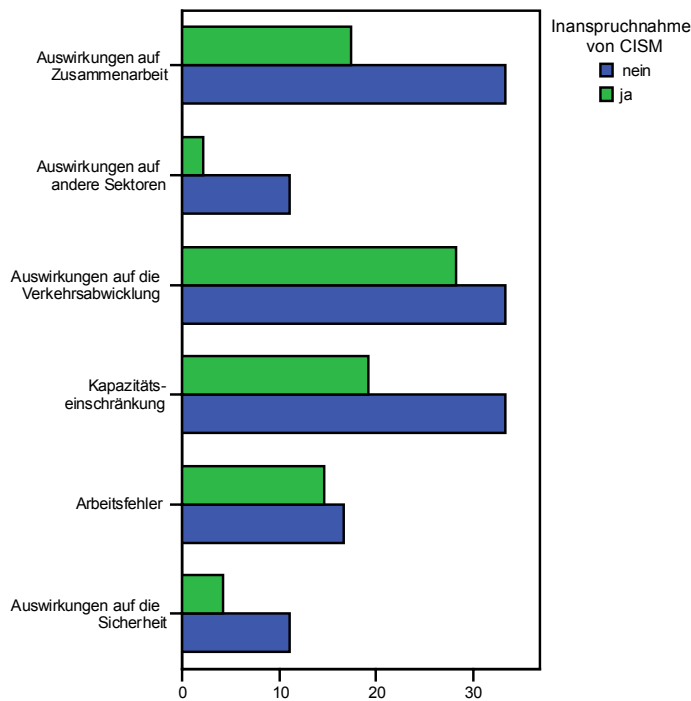


Abbildung 23: Auswirkungen der Beeinträchtigung nach kritischen Ereignissen auf den Flugsicherungsprozess

Resumé zur Selbststeuerung der Fluglotsen nach einem kritischen Ereignis

Die Fluglotsen fühlten sich durch das kritische Ereignis zum großen Teil erheblich belastet und eingeschränkt. Sie sahen aber gute eigene Möglichkeiten, diese zu bewältigen und wurden in ihren Bewältigungsressourcen durch CISM unterstützt. Fluglotsen, die dies nutzten, schilderten sich als weniger eingeschränkt in fundamentalen, aufgabenbezogenen Fähigkeiten und in den planerischen und interaktiven Kerntätigkeiten des Fluglotsen. Dies spiegelte sich dann darin wider, dass sie in Bezug auf die Arbeitsgüte und die Arbeitsmenge geringere Einbußen berichteten als die Fluglotsen ohne CISM. Auffällig war auch die größere, selbstkritische Bewusstheit der CISM-Fluglotsen für ihre Belastungsfähigkeit. Insofern fördert CISM die Selbststeuerung des Fluglotsen im Sinne der Unternehmensziele. Fluglotsen ohne CISM-Gespräch reagierten auf ihre kognitiven Einschränkungen mit einer Verminderung der Zusammenarbeit und größerer Entschlossenheit im Handeln. Mit dieser Entscheidungsstärke war allerdings eine weniger umsichtige und kooperative Planung und Steuerung verbunden, unter der sowohl der Verkehrsfluss als auch die Produktivität litten, weil beide auch auf guter Teamarbeit basieren.

2.2.7.2 Einfluss der operativen Steuerung

Der zweite Auswertungsbereich betraf die Frage, wie die operative Führung (Prozesssteuerung) mit dem kritischen Ereignis und der Beeinträchtigung des Fluglotsen umgeht. Hier wurden zwei Aspekte geprüft:

1. Welche Maßnahmen hat die Führungskraft unterstützt bzw. initiiert?

2. Wie haben sich diese auf die Leistungsfähigkeit des Fluglotsen und die Funktionsfähigkeit des ATC-Systems ausgewirkt?

Welche Maßnahmen hat die Führungskraft unterstützt bzw. initiiert?

Die Interviews mit den Führungskräften ergaben, dass einer Führungskraft nach einem kritischen Ereignis verschiedene Interventionsmöglichkeiten zur Verfügung stehen: Das persönliche Gespräch mit dem Fluglotsen, die Ablösung aus dem Dienst am gleichen Tag, die Anordnung eines Sektorwechsels, die Ablösung für eine Sonderaufgabe, die Verminderung der Arbeitslast (Sektorsteuerung, s. Kapitel 2.2.5.1) sowie das Initiieren des CISM-Gespräches. Die Fluglotsen wurden entsprechend gefragt, ob es nach dem Ereignis Veränderungen im Einsatzplan gab, wer das CISM-Gespräch initiiert hat oder ob es Veränderungen der Kapazität/Arbeitslast im eigenen Sektor gab.

Die Daten zu diesen verschiedenen Fragen und die offenen Antworten zu der Frage nach den Reaktionen anderer auf das kritische Ereignis wurden für jeden Fluglotsen zusammengestellt und dann kategorisiert. Es zeigte sich, dass nur in je einem Fall ein Sektorwechsel oder eine Sektorsteuerung vorgenommen wurden. Auch die Anordnung einer Sonderaufgabe wurde nur in einem Fall beschrieben. Um statistisch auswertbare Gruppen zu erhalten, wurden daher lediglich gehäufte Interventionsformen betrachtet, mit denen verwertbare Trends zur Wirkung dieser Eingriffe beschrieben werden konnten.

Die häufigste Interventionsform nach „Initiieren von CISM“ war „Ablösung aus dem Dienst am gleichen Tag“ (mit oder ohne Sonderaufgabe). Des Weiteren gab es nur in zwei Fällen ein Führungsgespräch mit dem Fluglotsen. Da einige Fluglotsen zusätzlich die Gespräche mit Kollegen als hilfreich nannten, wurden diese beiden Formen der kommunikativen Unterstützung zu einer Interventionsform zusammengefasst. Als weitere Variable wurden der Zeitpunkt und die Häufigkeit von CISM-Gesprächen in diese Analyse aufgenommen, da beides entscheidend vom Führungshandeln abhängt und nach dem Verständnis der Krisenintervention relevant für die Erholung ist. Somit ergaben sich sechs Gruppen mit folgenden Häufigkeiten (s. Abbildung 24):

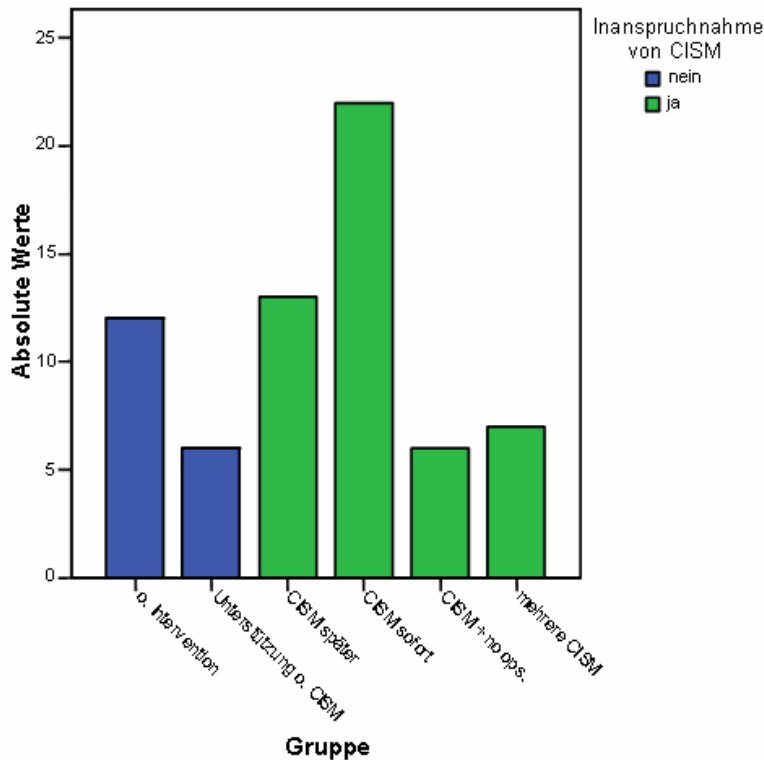


Abbildung 24: Die verschiedenen Interventionsformen der Prozesssteuerung nach einem kritischen Ereignis

Die Gruppen repräsentieren folgende Interventionsformen (von links nach rechts):

Gruppe 1.: Ohne Intervention. Kein Eingriff und kein CISM-Gespräch (in 12 Fällen)

Gruppe 2.: Intervention ohne CISM. Kein CISM, keine Eingriffe aber Unterstützung durch Kollegen oder durch Gespräche mit dem Vorgesetzten (in 6 Fällen)

Gruppe 3.: CISM später. Durchführung eines CISM-Gesprächs an einem anderen Tag ohne sofortigen Eingriff (in 13 Fällen)

Gruppe 4.: CISM sofort. Sofortige Ablösung für ein CISM-Gespräch und Wiederaufnahme der Tätigkeit danach (in 22 Fällen)

Gruppe 5.: CISM *no ops.* Sofortige Ablösung für ein CISM-Gespräch und Ablösung aus der operativen Tätigkeit in der gleichen Schicht (in 6 Fällen)

Gruppe 6.: Mehrere CISM. Sofortige Ablösung für ein CISM-Gespräch und Durchführung weiterer CISM-Gespräche an den folgenden Tagen (in 7 Fällen)

Die Art des Führungshandelns scheint dabei unabhängig von der direkten emotionalen Betroffenheit der Fluglotsen zu sein, die in allen Gruppen bei etwa 82% lag (Kapitel 2.2.7.1).

Wie wirkten sich diese Führungseingriffe auf die Leistungsfähigkeit des Fluglotsen und die Funktionsfähigkeit des ATC-Systems aus?

Zunächst wurde die Auswirkung auf die Leistungsfähigkeit anhand der „Performance-Kurve“ analysiert. Die Differenzierung der Leistungskurven für diese sechs Gruppen zeigt Abbildung 25.

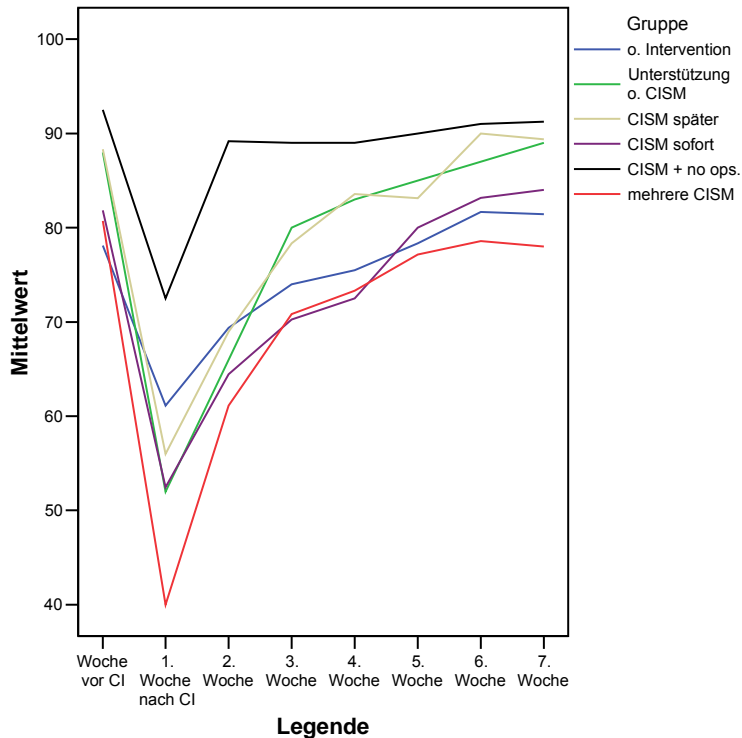


Abbildung 25: Performance-Kurven nach Interventionsformen (CI: Critical Incident)

Auf den ersten Blick zeigten alle Kurven einen vergleichbaren Verlauf allerdings auf unterschiedlichem Niveau. Insbesondere drei Gruppen fielen durch einen abweichenden Verlauf von der Durchschnittskurve auf: Die **Gruppe 6** (rote Kurve) hatte den stärksten subjektiven Leistungseinbruch und die längste Zeit der Beeinträchtigung. Die Fluglotsen nahmen hier mehrere CISM-Gespräche in Anspruch. Die **Gruppe 5** (schwarze Kurve) hatte den geringsten subjektiven Leistungseinbruch und die kürzeste Zeit der Beeinträchtigung. Die Daten legen also nahe, dass die Nutzung eines CISM-Gesprächs und die Ablösung aus dem operativen Dienst am gleichen Tag die Erholung in besonderem Maße fördert. Die **Gruppe 1** (blaue Kurve) zeigte nach einem „normalen“ Leistungseinbruch eine vergleichsweise flache Erholungskurve. Ohne jegliche Unterstützung wurde die Erholung offensichtlich nicht in gleichem Maße gefördert wie in den anderen Gruppen. Die **Gruppen 2 und 3** (grüne und gelbe Kurve) wiesen einen fast identischen Verlauf auf, d. h. verspätetes CISM wirkt ebenso wenig wie kollegiale Unterstützung ohne CISM. Die Differenzierung der Kurven zeigt auch, dass der Vergleich der „CISM-Kurve“ mit der „Nicht-CISM-Kurve“ zunächst irreführend sein kann, da diese beiden Gruppen in sich sehr heterogen sind.

Eine Überprüfung der Ursachen für diese Unterschiede in der generellen Leistungsfähigkeit sollte dann in einem Vergleich der Beeinträchtigungen und der Tätigkeitse-

inschränkungen für diese sechs Gruppen vorgenommen werden. Das Profil der wichtigsten Tätigkeiten im Gruppenvergleich zeigt Abbildung 26.

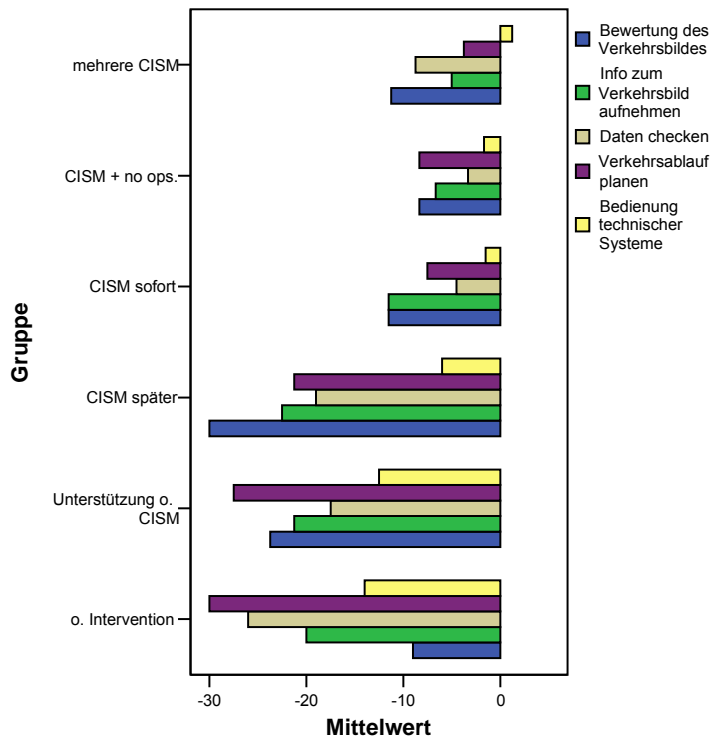


Abbildung 26: Tätigkeitseinschränkungen in den sechs Interventionsgruppen

Es fällt auf, dass drei Gruppen erheblich mehr Einschränkungen in der Tätigkeitsausübung erlebten als die anderen drei Gruppen. Neben den beiden Gruppen ohne CISM-Unterstützung (Gruppe 1 + 2) war dies auch die Gruppe 3, die ein CISM-Gespräch zu einem späteren Zeitpunkt in Anspruch nahm bzw. angeboten bekam. Es ist zu vermuten, dass das CISM-Gespräch und die damit verbundene kognitive Aufarbeitung des Ereignisses „den Kopf frei macht“ für die anschließende Lotsentätigkeit. Ohne dieses Gespräch arbeiteten die Fluglotsen das Ereignis möglicherweise während ihrer operativen Tätigkeit auf. Dies wirkte sich natürlich auf die Tätigkeitsausübung aus, da diese die volle kognitive Kapazität erfordert. Insbesondere in der Gruppe ohne jegliche Hilfestellung fanden sich starke Einschränkungen in der „Datenverarbeitung“. Aber auch mit kollegialer Unterstützung allein bzw. dem Gespräch mit dem Vorgesetzten (Gruppe 2) oder einem verspäteten CISM (Gruppe 3) fanden sich ähnliche Einschränkungen in der Tätigkeitsausübung. Soziale Unterstützung allein oder verspätetes CISM scheint im Sinne der Wiederherstellung der Arbeitsfähigkeit kein Ersatz für professionelles CISM zu sein. Diese Interpretation wird gestützt durch die Einschätzungen der erlebten emotionalen und körperlichen Beeinträchtigungen für diese Gruppen. Wer keine Unterstützung durch CISM erfährt, neigt im Belastungszeitraum eher zu emotionalen und körperlichen Dysfunktionen, wie Schlafstörungen, Gereiztheit, Schuldgefühle und negativen Einstellungen, auch wenn kollegiale Unterstützung erfahren und diese sehr positiv geschildert wird.

Bei der Gruppe 6 (mehrere CISM-Gespräche) war kein Zusammenhang zwischen der stark erlebten generellen Leistungsbeeinträchtigung und den Defiziten in der Tätigkeitsausübung zu erkennen. Möglicherweise ist das Gefühl der Leistungsfähigkeit

unabhängig von der realen kognitiven Kapazität für die Tätigkeitsausübung und wird durch die mehrfachen CISM-Interventionen gefördert. In diesem Fall der Gruppe 6 ist die *Performance*-Kurve damit kein sicherer Prädiktor für die Qualität der Tätigkeitsausübung.

Abschließend stellte sich die Frage, welche Auswirkungen die verschiedenen Interventionsformen auf die Leistungsfähigkeit des ATC-Systems haben. Dazu wurden die drei Ergebnisauswirkungen im Gruppenvergleich analysiert, die sich in der bisherigen Auswertung als beeinträchtigt gezeigt hatten. Abbildung 27 zeigt die Ergebnisse im Gruppenvergleich.

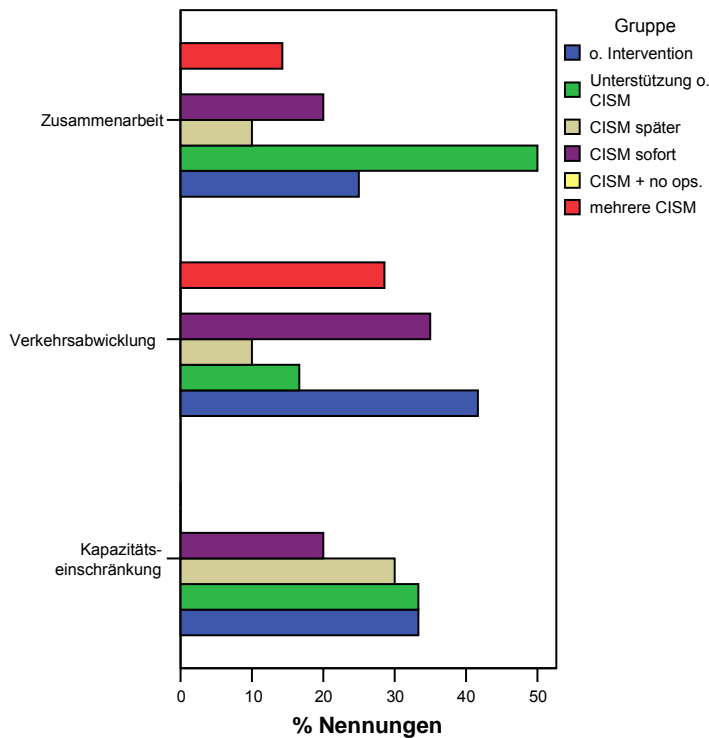


Abbildung 27: Ergebnisauswirkungen der verschiedenen Interventionsformen

Auch hier hob sich erneut die Gruppe 5 positiv von den anderen Gruppen ab. Von den Fluglotsen mit CISM und Ausplanung am gleichen Tag wurden überhaupt keine negativen Auswirkungen auf den Flugsicherungsprozess beschrieben (alle Balken auf 0 in Abbildung 27). Dies spricht dafür, dass diese „Wirkstoffkombination“ in der Lage ist, negative Effekte auf die Leistungsmenge und Leistungsgüte weitestgehend zu vermeiden.

Die deutlichsten Auswirkungen auf die Kapazität wurden analog zur beschriebenen Tätigkeitseinschränkung in den beiden „Nicht-CISM-Gruppen“ sowie in der Gruppe mit einem CISM-Gespräch zu einem späteren Zeitpunkt genannt. Die Häufigkeit genannter Auswirkungen auf die Zusammenarbeit in Gruppe 2 (Unterstützung ohne CISM) fällt auf. In dieser Gruppe wurden allerdings auch viele positive Teameffekte beschrieben, da hier konkrete kollegiale Unterstützung erlebt wurde. Daher könnte die Aufsummierung positiver und negativer Wirkungen die Häufigkeit der Nennung von Zusammenarbeitsauswirkungen verursacht haben.

Resuméé zur operativen Steuerung nach einem kritischen Ereignis

Der Eingriff der Führungskraft in die operative Steuerung war bedeutsam für die Erholung und Stabilisierung des Flugsicherungsprozesses. Insbesondere die Kombination eines sofortigen CISM-Gesprüches mit der Ablösung des betroffenen Fluglotsen aus der operativen Arbeit für den gleichen Tag bzw. die gleiche Schicht zeigte deutlich die geringsten Einschränkungen in der Tätigkeitsausübung und keine wahrgenommenen Einbußen in der Verkehrsabwicklung oder der bewältigten Arbeitslast. Es spricht vieles dafür, dass die Bewältigung hier durch die Kombination von sofortigem CISM mit der Führungsintervention „dienstfrei für den Rest der Schicht“ forciert wird und am nächsten Tag die kognitive Kapazität wieder für die operative Tätigkeit zur Verfügung steht.

2.2.7.3 Bedeutung für die strategische Steuerung

Die Wirtschaftlichkeitsanalyse im HPM (Kapitel 1.4, Abbildung 1) schließt ab mit der ökonomischen Bewertung der aufgezeigten Effekte mit Blick auf die strategischen Unternehmensziele und „*Key Performance Indicators*“ (EUROCONTROL, 2004a). Hier wurden folgende Fragen geprüft:

1. Wie lassen sich die aufgezeigten Wirkungen von CISM monetär bewerten?
2. Wie schneidet CISM ab, wenn man Kosten und Nutzeneffekte ins Verhältnis setzt?
3. Welche Bedeutung hat CISM mit Blick auf die strategischen Ziele und die *Key Performance Indicators*?

Wie lassen sich die aufgezeigten Wirkungen von CISM monetär bewerten?

Der monetäre Effekt von CISM lässt sich durch die Leistungskurve und die angegebenen Erholungsanteile innerhalb der CISM-Gruppe analog zur Machbarkeitsstudie (Kapitel 2.2.2) kalkulieren. Dies bietet zugleich die Möglichkeit, die Ergebnisse der Machbarkeitsstudie in Bezug auf diesen wichtigen und vielfach beachteten Punkt zu überprüfen.

Zu diesem Zweck wurde die Leistungskurve nur für die CISM-Gruppe berechnet (s. Abbildung 28). Sie unterscheidet sich nicht wesentlich von der *Performance*-Kurve für alle Fluglotsen. Dabei sei noch einmal erwähnt, dass aus der differenzierten Analyse der Gruppen deutlich wurde, dass diese Kurve in sich sehr verschiedene Formen und Intensitäten der Beeinträchtigung und Erholung enthält (Gruppe 3 – 6, Kapitel 2.2.7.2). Die folgende Berechnung ist daher konservativ, denn sie beruht auf allen CISM-Subgruppen. Würde man nur die *Best Practice* Subgruppe CISM *no ops* zugrunde legen, wären die monetären Effekte noch positiver.

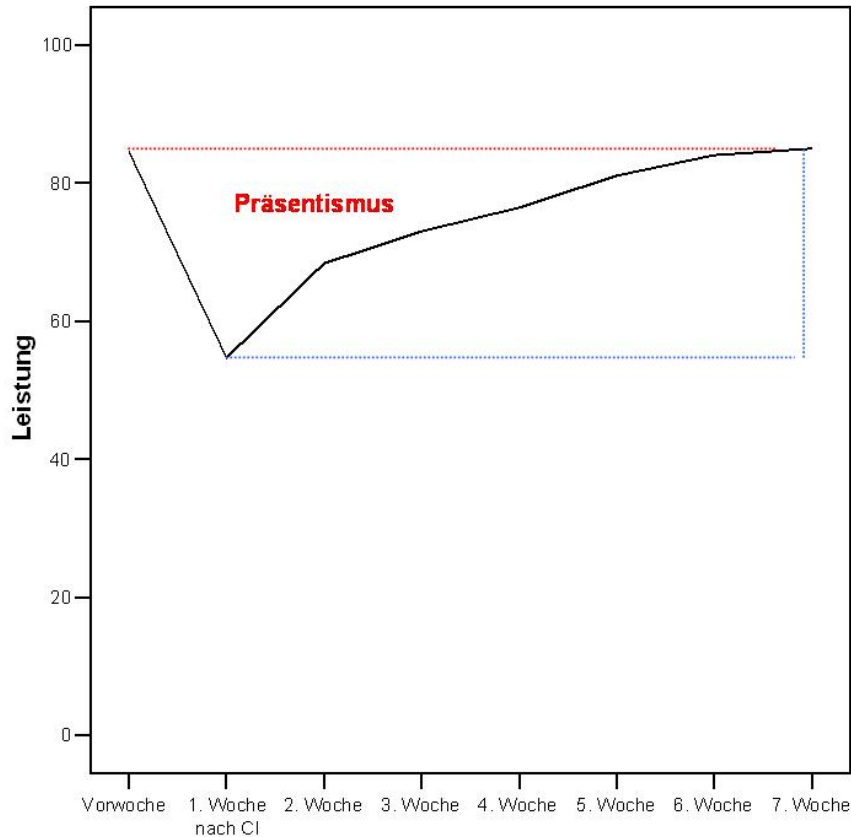


Abbildung 28: Performance-Kurve für die CISM-Gruppen (CI: Critical Incident)

Die Berechnung des monetär bewerteten Erholungseffektes erfolgte in gleicher Weise wie in der Machbarkeitstudie: Die durch die Erholung gewonnene Leistungsfähigkeit zwischen dem kritischen Ereignis und dem Endpunkt der Leistungskurve nach sieben Wochen entspricht für die CISM-Gruppe im Durchschnitt 8,5 Leistungstagen. An dieser gewonnenen Leistungszeit hat CISM wie bereits aufgezeigt einen Anteil von 36,3%, d.h. durch CISM wurden

$$8,5 * 0,363 = 3,0855 \text{ Tage}$$

Leistungszeit gewonnen. Das Kostenäquivalent dieser Leistungszeit ist

$$3,0855 * 670 = 2.067 \text{ €}$$

da die Vollkosten pro Lotsentag von der DFS mit 670 € beziffert werden.

Ohne CISM würden also 2.067 € kalkulatorische Kosten für zusätzlichen Leistungsausfall pro Fluglotse mit kritischem Ereignis anfallen. Für die untersuchte Stichprobe von 48 betroffenen Fluglotsen ergibt sich eine Kosteneinsparung von

$$48 * 2.067 = 99.216 \text{ €}.$$

Wie schneidet CISM ab, wenn man Kosten und Nutzeneffekte ins Verhältnis setzt?

Die Kosten für CISM betragen seit der Einführung im Jahre 1998 560.000 €. Darin enthalten sind Schulungskosten und Personalkosten für Abwesenheit der Fluglotsen während ihrer Aus- und Fortbildung als kollegiale Berater. Abwesenheitskosten für

die Gespräche selbst sind nicht eingerechnet, da diese den Gesamtbetrag nur marginal erhöhen. Die anteiligen Kosten für die untersuchte Stichprobe (16,7% der deutschen Fluglotsen) und den untersuchten Zeitraum von 28 Monaten (2,33 Jahre), bei einer Gesamtlauzeit des DFS CISM-Programms von 8 Jahren betragen:

$$560.000 \text{ €} / 8 \text{ Jahre} * 0,167 * 2,33 \text{ Jahre} = 27.238,00 \text{ €}$$

Der *Return on Investment* (ROI) lässt sich berechnen durch:

$$\text{ROI} = 99.216 / 27.238 * 100 = 364\%$$

Damit wird der ROI aus der Machbarkeitsstudie (Kapitel 2.2.2) von etwa 250% sogar noch übertroffen. Zurückzuführen ist dies vor allem auf den verlängerten Zeitraum, für den die *Performance*-Kurve und damit die Leistungseinschränkung erhoben wurden. Dies erhöht den gewonnen Leistungswert und damit den errechneten Nutzen von CISM.

Welche Bedeutung hat CISM mit Blick auf die strategischen Ziele und die *Key Performance Indicators*?

Ein wichtiges Ziel der Hauptstudie war es, die Kosten- und Nutzeneffekte von CISM auf Prozess- und Organisationsebene real zu erfassen. Der berechnete ROI würde jedoch nur dann die tatsächliche Kostenersparnis durch CISM widerspiegeln, wenn

1. die durch CISM gewonnene individuelle Leistungszeit 1:1 in eine verbesserte Systemleistung des Flugsicherungsprozesses (Team, benachbarte Sektoren, etc.) münden würde und
2. diese verbesserte Systemleistung direkt zu entsprechenden Personalkapazitätsanpassungen führen würde bzw. die freiwerdende Kapazität sinnvoll in andere Wertschöpfungsprozesse eingebracht werden könnte.

Die *Performance*-Kurve hat sich bewährt, um negative Leistungseffekte in der Folge kritischer Ereignisse und positive Leistungseffekte aufgrund des HF-Programms CISM und einer optimalen operativen Führung darzustellen. Daher wird die Erhebung einer *Performance*-Kurve auch für andere HF/HR/T-Evaluationsstudien empfohlen. Wie die komplexen Zusammenhänge zwischen der generellen Leistungsfähigkeitseinschätzung (*Performance*-Kurve) und den offen beschriebenen Tätigkeitseinschränkungen und Ergebnisauswirkungen jedoch zeigten, ist die *Performance*-Kurve nur ein Anhaltspunkt, wenn auch ein wichtiger, um die Bedeutung von gewonnener Leistungszeit bzw. vermiedenem Präsentismus zu schätzen. Spezifischere Indikatoren für die wirtschaftliche Bedeutung von CISM ergeben sich eher aus den aufgezeigten Ergebniswirkungen im Bereich der strategischen Ziele der DFS.

Die Unternehmens*performance* fokussiert auf drei zentrale Zielbereiche: Sicherheit, Produktivität und Pünktlichkeit. Die Auswirkungen von Beeinträchtigungen nach kritischen Ereignissen auf diese drei Leistungsaspekte wurden durch die betroffenen Fluglotsen in der Befragung eingeschätzt und in Kapitel 2.2.7.2 dargestellt.

Es wurde deutlich, dass insbesondere die Kapazität und Verkehrsabwicklung, in geringem Umfang auch die Sicherheit, durch CISM stabilisiert werden.

Sicherheit und Pünktlichkeit sind qualitative Ziele für die DFS, die nicht unmittelbar mit ökonomischen Größen verbunden sind. Probleme in der Pünktlichkeit, die durch den Prozess der Flugsicherung verursacht werden, z.B. spätere Startzeiten oder Abweichungen von der besten Streckenführung, führen nicht bei der DFS, sondern bei den Fluglinien zu hohen Kosten. Die Zufriedenheit der Fluglinien als Kunden der DFS kann dadurch erheblich belastet werden und im Kampf um Wettbewerbsvorteile im *Single European Sky* daher große Bedeutung haben. Fast jeder zweite Fluglotse, der keine Unterstützung nach einem kritischen Ereignis bekommen hat (Gruppe 1), berichtete von Einschränkungen in der Flüssigkeit der Verkehrsabwicklung in den zwei Wochen nach dem kritischen Ereignis. Auswirkungen auf die Pünktlichkeit wurden dabei z.T. explizit genannt. Insofern trägt CISM erheblich zur Erhöhung der Kundenzufriedenheit bei, da es die Faktoren Sicherheit und Pünktlichkeit fördert.

Für den Zielbereich Kapazität lagen auch quantitative Daten vor, da die Fluglotsen gebeten wurden, soweit möglich, auch den Umfang und die Dauer der Ergebnisauswirkung der eigenen Beeinträchtigung für jedes Ergebnisziel anzugeben. Aus den Angaben der Kapazitätseinschränkung auf einer Prozentskala und der angegebenen Dauer der Einschränkung in Tagen ließ sich ein Gesamtwert der Leistungseinbuße berechnen. In den beiden Gruppen ohne CISM-Gespräch ergab sich im Durchschnitt ein Kapazitätsverlust von 77%. Dies entspricht einer verringerten Flugsicherungsleistung von 10% an 7,7 Tagen für jeden von einem kritischen Ereignis betroffenen Fluglotsen, der kein CISM-Gespräch in Anspruch nimmt.

Dieser Leistungsverlust wirkt sich kalkulatorisch auch auf die Gesamtproduktivität der DFS aus. Diese Zusammenhänge zwischen individueller Leistung und Produktivität des Unternehmens ergeben sich folgendermaßen: Werden Leistungseingänge durch einen Fluglotsen in einem Sektor verursacht, ist ein Nachbarsektor gleichermaßen betroffen. Die Leistung im Flugsicherungsprozess wird in der DFS für eine so genannte Einsatzberechtigungsgruppe (EBG) berechnet, bei der z.B. vier Sektoren zu einer Leistungseinheit zusammengefasst sind. Die Produktivität berechnet sich aus dem Verhältnis von bewältigten Flugstunden (Performance) und der Anzahl von Fluglotsen in der operativen Arbeit. Eine Leistungsminderung in einem Sektor von 10% entspricht dann einem Verlust an Kapazität von 5% in der betroffenen EBG. In der DFS arbeiten 26 EBGen. Der Gesamtleistungsverlust nach einem kritischen Ereignis beträgt somit 0,2% ($5\% / 26$) an 7,7 Tagen. Pro Tag werden in der DFS ca. 780.000 Flugminuten gesichert. Der Leistungsverlust pro Ereignis sind somit etwa 1.560 Flugminuten (0,2% von 780.000) und bei 100 Ereignissen pro Jahr 156.000. Flugminuten. Bei konstanter Personalkapazität in der operativen Arbeit entspricht der Leistungsverlust einem gleich großen Produktivitätsverlust von 0,055% ($= 156.000. / (780.000 \times 365 \text{ Tage})$). Für eine monetäre Bewertung lässt sich das Dienstleistungsentgelt einer Flugminute heranziehen, das von den Airlines vergütet wird. Dies sind zurzeit durchschnittlich ca. 30€ pro Flugminute. Damit ergibt sich eine Erlösverbesserung pro Ereignis von 46.800€ wenn ein CISM-Gespräch durchgeführt wird und eine Ablösung aus dem operativen Dienst erfolgt. Für 100 Ereignisse sind es entsprechend 4,68 Mio. €. Selbst wenn die Kapazitäts- und Erlöswirksamkeit von CISM erheblich überschätzt würde, zeigt dies, wie erlösensibel Kapazitätsveränderungen sind.

Dagegen sind die zusätzlichen Personalkosten, die durch die Ablösung entstehen, im Durchschnitt pro Ereignis ein halber Lotsentag, äußerst gering. Wenn diese zusätzlichen, nicht produktiven Lotsentage in die Personalplanung eingehen, müssen für 100 kritische Ereignisse 50 Lotsentage Kapazitäten zusätzlich eingeplant werden.

Dies entspricht Personalkosten von 33.500€. Selbst wenn die Kapazitätsverminderung nach dem kritischen Ereignis deutlich niedriger liegt, beispielsweise nur 0,1% an 7,7 Tagen, läge die Erlösverbesserung immer noch über den Personalkosten für die Ablösung.

Wenn auch die hier dargestellten Bezüge nur Trends darstellen und die Kalkulationen teilweise auf Schätzungen zurückgreifen mussten, so legen die Angaben und Beschreibungen der betroffenen Fluglotsen nahe, dass durch CISM auch strategisch bedeutsame Folgewirkungen der Beeinträchtigung des Fluglotsen vermieden werden können (erhebliche Produktivitätsnachteile und Verspätungen) und der Deckungsbeitrag der Niederlassungen positiv beeinflusst wird. In den vorangegangenen Abschnitten wurde nur beispielhaft demonstriert, welches Ausmaß die ökonomischen Effekte annehmen können. Es wurde in Anbetracht des Umfangs, den die CISM-Evaluation bereits eingenommen hat, darauf verzichtet, etwa mit Hilfe einer Szenarien-Technik die Spannweite möglicher Effekte aufzuzeigen. Dies wäre jedoch für zukünftige ökonomische Evaluationen ein empfehlenswertes Vorgehen.

2.2.7.4 Langfristige Folgen von kritischen Ereignissen

Bisher wurde in der Wirtschaftlichkeitsberechnung davon ausgegangen, dass kritische Ereignisse ausschließlich zu Folgewirkungen in den Wochen nach dem Ereignis führen. Es ist aber bekannt, dass kritische Ereignisse durch so genannte *déjà-vus* lange nachwirken können, insbesondere wenn sie nicht – wie es CISM explizit vorsieht – zeitnah behandelt werden. Im Abschnitt 4 des Fragebogens wurde erhoben, in welchem Umfang solche Nachwirkungen bei Fluglotsen auftreten. Dazu wurden die Fluglotsen nach ihrem bisher kritischsten Ereignis aus ihrer Arbeit als Fluglotse gefragt. Es sollte länger als zwei Jahre zurückliegen, um Doppelungen mit dem Abschnitt 2 auszuschließen und langfristig wirksame Ereignisse zu erfassen.

Insgesamt schilderten 149 Fluglotsen (49% der Stichprobe) ein länger als zwei Jahre zurückliegendes kritisches Ereignis. 47% der Fluglotsen gaben dabei an, in den folgenden Jahren *déjà-vu*-Situationen erlebt zu haben. Die Abbildung 29 zeigt an, wie häufig diese *déjà-vus* von den Fluglotsen erlebt wurden:

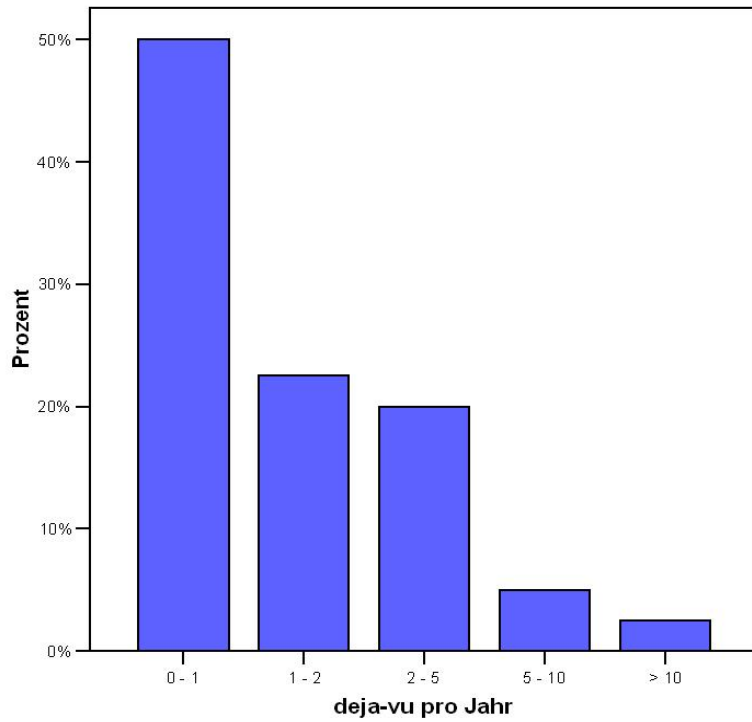


Abbildung 29: Häufigkeit von *déjà-vus* bei Fluglotsen mit mehr als zwei Jahre zurückliegendem kritischem Ereignis
Bei 50% der Fluglotsen tauchten *déjà-vus* mehr als einmal pro Jahr wieder auf. Dabei stellte sich die Frage, welche Konsequenzen dies auf die Befindlichkeit und die Leistungsfähigkeit des Fluglotsen hat. Die Auswirkungen auf die Befindlichkeit zeigt Abbildung 30.

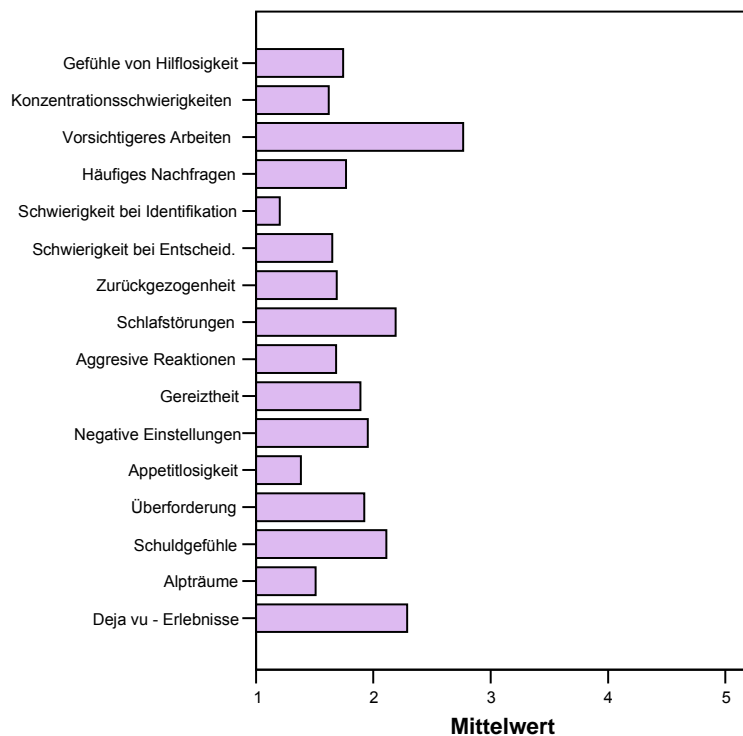


Abbildung 30: Beeinträchtigungen durch *déjà-vu*-Situationen (1 nicht, 2 kaum, 3 einigermaßen, 4 erheblich, 5 stark aufgetretene Beeinträchtigungen)

Das Beeinträchtigungsprofil zeigt große Ähnlichkeiten mit dem Profil nach einem aktuellen oder weniger als zwei Jahre zurückliegenden kritischen Ereignis, wobei die Symptome etwas weniger stark ausgeprägt sind. Auch hier dominierten „vorsichtigeres Arbeiten“, „Schlafstörungen“ und „Schuldgefühle“. Dies bedeutet, dass nach kritischen Ereignissen mit starker emotionaler Belastung die in Kapitel 2.2.7.1 geschilderten Konsequenzen Jahr für Jahr bei der Hälfte der Fluglotsen mehrfach wieder auftauchen.

Dass dies auch Konsequenzen für die Leistungsfähigkeit der Flugsicherung hat, zeigt der für *déjà-vu*-Situationen berechnete Präsentismuseffekt. Aus den Angaben der Fluglotsen zur Leistungsminderung bei *déjà-vus* konnte auch hier ein *Performance*-Verlustwert ermittelt werden. Im Durchschnitt erlebte ein Fluglotse mit *déjà-vus* jedes Jahr eine Leistungseinschränkung von insgesamt 1,87 Arbeitstagen. Der gemittelte Wert ergibt sich aus ca. 2,5 Tagen mit leichter, ca. 1,8 Tagen mit mittlerer und ca. ½ Tag mit starker Leistungseinschränkung pro Jahr. Dies entspricht etwa 35% des Präsentismuseffekts direkt nach kritischen Ereignissen (Kapitel 2.2.7.3). Der Anspruch von CISM ist, genau diese *déjà-vus* aufgrund fehlender oder unzureichender Verarbeitung des Ereignisses weitestgehend zu vermeiden. Eine Untersuchung, ob CISM hier präventiv wirkte, konnte leider nicht durchgeführt werden, weil viele geschilderte Ereignisse sich in Jahren ereigneten, in denen das CISM-Programm der DFS noch nicht implementiert war. Es wird aber deutlich, dass Folgen von kritischen Ereignissen noch viele Jahre nach dem kritischen Ereignis auftreten können und dann erneut Konsequenzen für die Sicherheit, Pünktlichkeit und Produktivität haben können. Der aufgezeigt Nutzenbeitrag von CISM ist damit wahrscheinlich noch wesentlich höher anzusetzen als in Kapitel 2.2.7.3 konservativ kalkuliert. Literaturanalysen haben ge-

zeigt, dass CISM entwickelt wurde, um solche *déjà-vus* zu vermeiden und in diesem Sinne erfolgreich ist. Die Wirksamkeit von CISM, wenn es den Standards entsprechend eingesetzt wird, wurde in 65 Studien nachgewiesen (EUROCONTROL, 2004b).

2.2.7.5 Wechselwirkungen von CISM und Sicherheitskultur

In der Machbarkeitsstudie (Kapitel 2.2.2) ergaben sich Hinweise auf eine Verbesserung der Kommunikations- und Sicherheitskultur bei der DFS durch die Einführung des CISM-Programms. Diese intangiblen Nutzeneffekte sollten in der großen Stichprobe der Hauptstudie weiter untersucht werden. Abschnitt 3 des verwendeten Fragebogens erfasste daher die Kommunikations- und Sicherheitskultur bei der DFS im Allgemeinen und ihre Veränderung durch CISM im Besonderen. Zur Kontrolle wurde auch erfragt, welche unternehmenskulturellen Beiträge anderen HF/HR/T-Programmen zugeschrieben werden.

Die Ergebnisse zu den drei wesentlichen Fragen dieses Themenkomplexes werden im Folgenden dargestellt:

1. Wie wird die Kommunikationskultur bei der DFS allgemein beschrieben?
2. Wie hat CISM diese Kommunikationskultur verändert?
3. Wie ist die kulturverändernde Wirkung im Vergleich zu den Effekten anderer Programme?

Abbildung 31 zeigt das Kulturprofil der DFS. Die Befragten gaben an, inwiefern kulturbeschreibende Adjektive auf die Kultur bei der DFS zutreffen. Die durchschnittlichen Bewertungen bewegen sich zwischen 3 und 4, also im Mittelfeld der Skala.

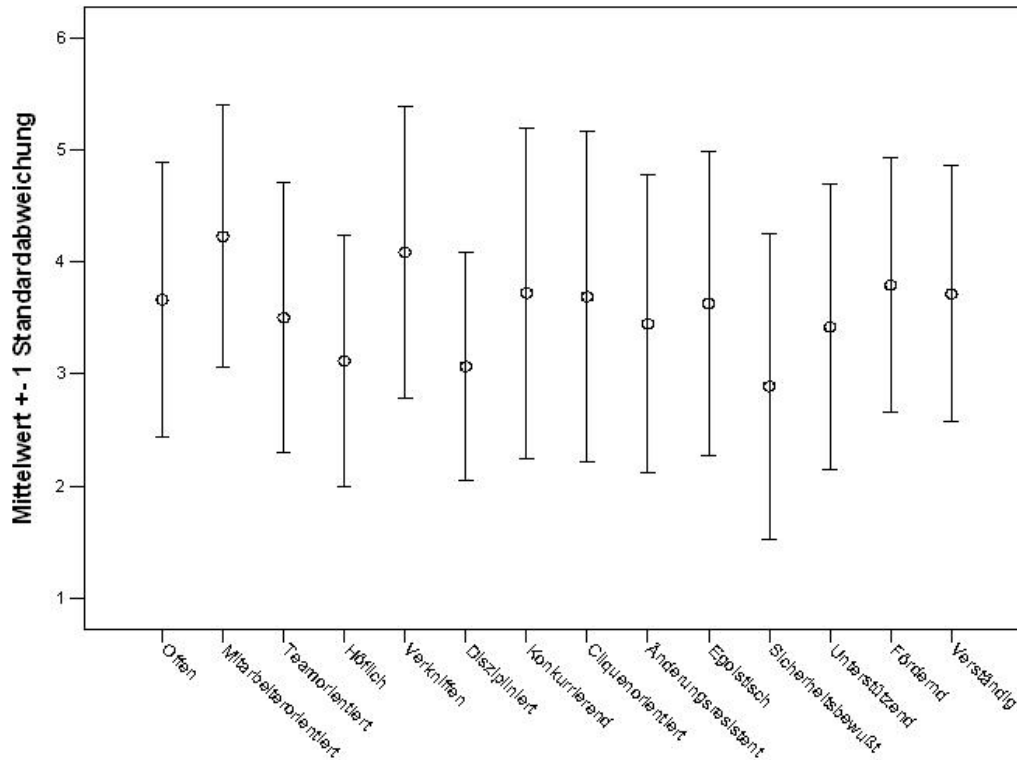


Abbildung 31: Kulturprofil der DFS anhand von Adjektiven (1: trifft genau zu, 6: trifft nicht zu)

Die höchste Zustimmung fand erwartungsgemäß das Adjektiv „sicherheitsorientiert“. Als für die DFS am wenigsten zutreffend empfanden die Fluglotsen das Adjektiv „mitarbeiterorientiert“. An den großen Standardabweichungen ist abzulesen, dass sich die Befragten sehr uneinig waren.

Abbildung 32 zeigt die Veränderung dieses Kulturprofils durch das CISM-Programm. Hier herrscht größere Einigkeit unter den Befragten: Hinsichtlich aller Adjektive gaben sie mit vergleichsweise kleinen Standardabweichungen an, dass CISM die Kultur bei der DFS verbessert hat. Die mittlere Verbesserung erreichte zwar nur für das Adjektiv „unterstützend“ +1, statistisch gegen 0 getestet sind jedoch fast alle Verbesserungen signifikant (Irrtumswahrscheinlichkeit $p < 0,05$). Nur die Eigenschaften „verkniffen“, „konkurrierend“ und „cliquenorientiert“ konnte CISM nicht statistisch signifikant verbessern. Strategisch bedeutsam hinsichtlich der Unternehmensziele der DFS ist vor allem die signifikante Verbesserung der Sicherheitsorientierung (viertletzte Adjektiv in Abbildung 32).

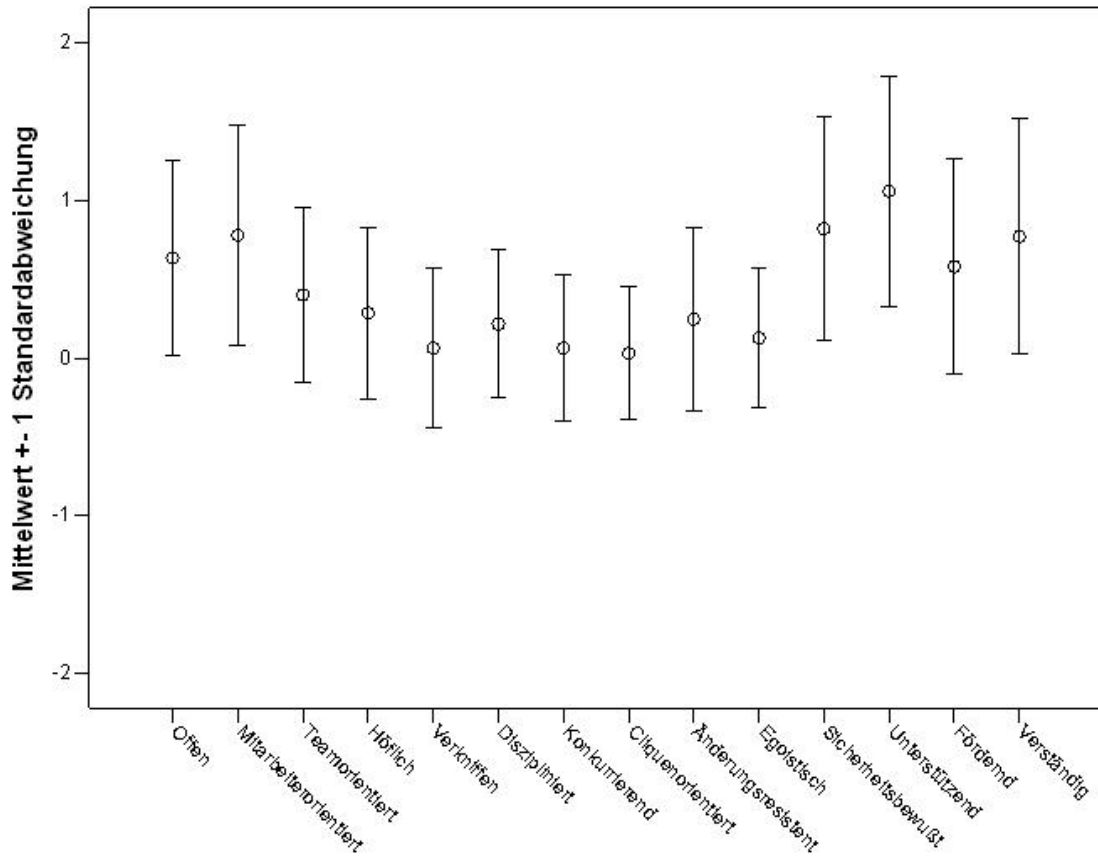


Abbildung 32: Veränderung des Kulturprofils der DFS durch CISM (+2 stark verbessert, -2 stark verschlechtert)

Abbildung 33 zeigt die prozentuale Veränderung der Sicherheitskultur bei der DFS durch verschiedene HF/HR/T-Programme. Die Befragten konnten sowohl negative (Verschlechterung) als auch positive Prozentangaben (Verbesserung) machen. Ein Drittel der positiven Veränderung der Sicherheitskultur und damit der größte Beitrag wurde CISM zugeschrieben. Alle anderen HF/HR/T-Programme leisteten nur Beiträge bis zu 20%.

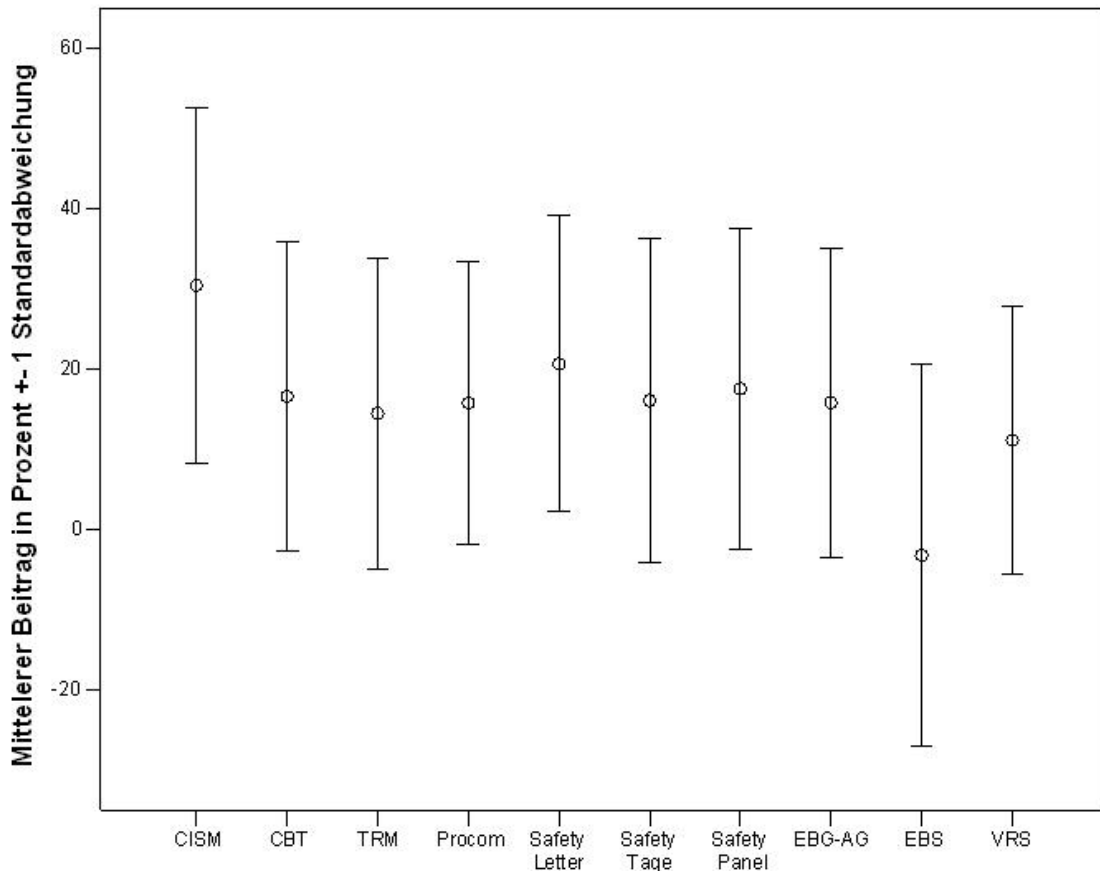


Abbildung 33: Beitrag verschiedener HF/HR/T-Programme zur Veränderung der Sicherheitskultur der DFS

| | |
|----------------------|---|
| CBT | <i>Computer Based Training</i> für Flugsicherungsverfahren seit 1999 |
| TRM | <i>Team Resource Management</i> für Kommunikation/Kooperation seit 1999 |
| Procom | Professionalität und Kommunikation; Beobachtung und <i>Feedback</i> seit 2000 |
| <i>Safety Letter</i> | Vierteljährliche Information über kritische Ereignisse seit 1999 |
| <i>Safety Tage</i> | Mindestens jährliche <i>Briefings</i> zu sicherheitsrelevanten Themen |
| EBS | <i>Electronic Briefing System</i> , in dem Fluglotsen nach Arbeitsabwesenheit aktuelle Informationen lesen müssen |
| VRS | <i>Voluntary Reporting System</i> , Fluglotsen berichten über kritische Ereignisse |

Zusammenfassend wurde das vorläufige Ergebnis der Machbarkeitsstudie (Kapitel 2.2.2), dass CISM weit reichende unternehmenskulturelle Nutzeneffekte auslöste, in vollem Umfang bestätigt. Das CISM-Programm übertrifft in dieser Hinsicht alle anderen HF/HR/T-Programme der DFS. Die oben beschriebenen monetären Nutzeneffekte des *Return on Investment* und der Kapazitätsverlustprävention sind auch aufgrund der hier gefundenen intangiblen Kulturverbesserungen Unterschätzungen des wirklichen Werts des CISM-Programms für die DFS.

Die aus den verschiedenen Abschnitten gewonnen Erkenntnisse können abschließend in Form einer Wirkungskette zusammengefasst werden (s. Abbildung 34).

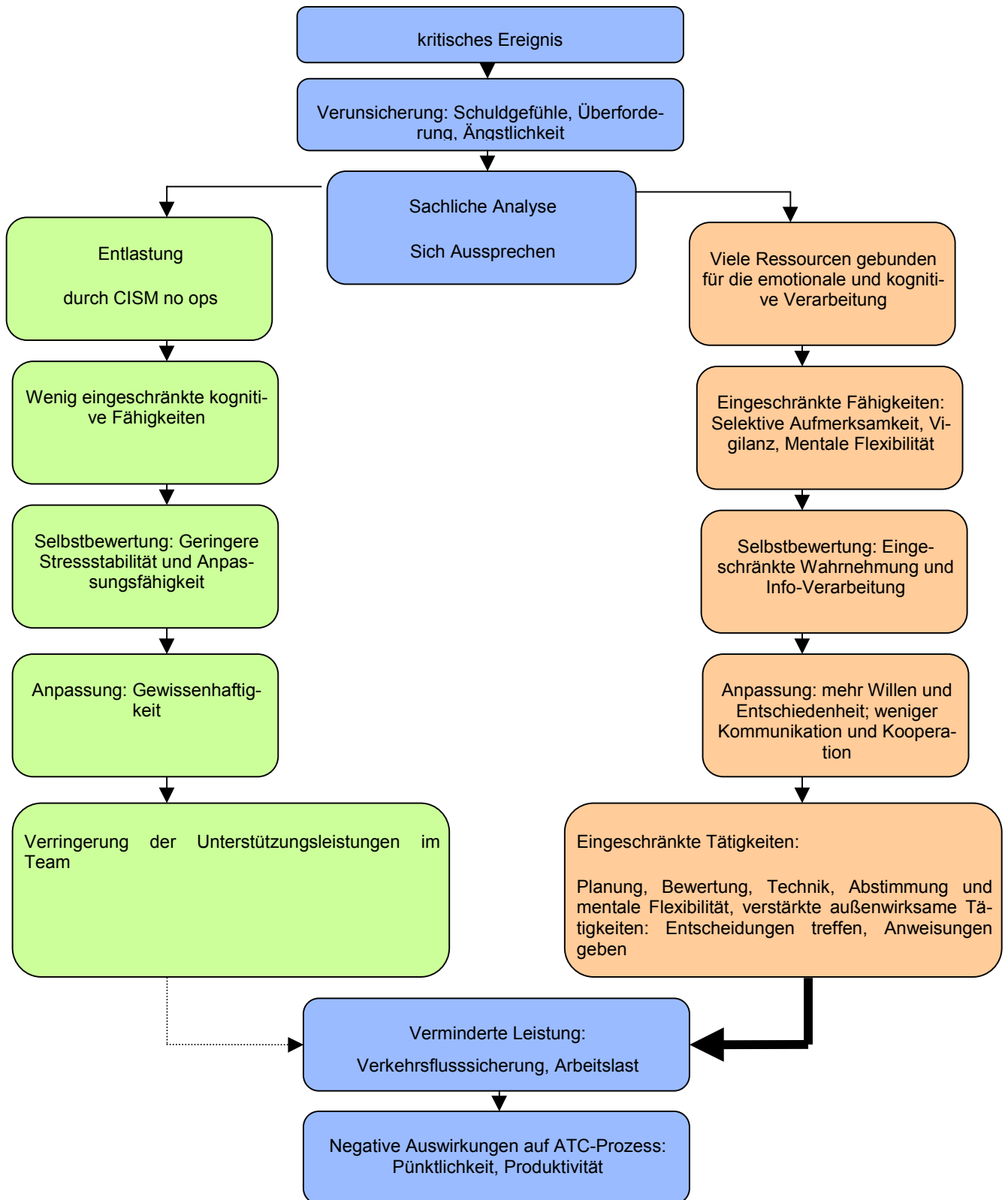


Abbildung 34: Wirkungskette der Reaktionen auf ein kritisches Ereignis (mit CISM links in grün und ohne CISM rechts in rot)

Die linke Seite zeigt die Verarbeitung des kritischen Ereignisses und die leistungsbezogenen Konsequenzen, wenn ein CISM-Gespräch in Anspruch genommen wird und danach eine Ablösung aus der operativen Tätigkeit für den Rest der Schicht vorgenommen wird. Die rechte Seite zeigt die Wirkungen ohne CISM. Mit diesem Verlauf werden die Hypothesen aus der Machbarkeitsstudie (Kapitel 2.2.2) bestätigt und weiter konkretisiert. Insbesondere wird deutlich, welche erfolgskritischen Wirkungen die Verarbeitung des Vorfalls durch den betroffenen Fluglotsen bei der Arbeit und ohne professionelle Hilfe hat. Sie nimmt kognitive Ressourcen in Anspruch, die für die Flugsicherungstätigkeit fehlen und in einer geringeren Produktivität münden.

2.2.8 Resumée zur CISM-Studie

Die bis hierher beschriebene Evaluation des HF-Programms CISM wird von den Autoren hinsichtlich Quantität und Qualität der Datenerhebung als Benchmark für die Praxis angesehen. Insbesondere konnte eine Wirkungskette aufgezeigt werden, die Funktionsebene (Fluglotse) mit Prozessebene (operative Führung, Flugsicherungsprozess) und Organisationsebene (strategische Führung) plausibel miteinander verbindet. Außer den subjektiven Interview- und Fragebogendaten aus Selbst- und Fremdsicht flossen Kennzahlen aus dem *Controlling*, dem *Safety*- und dem Qualitätsmanagement der DFS mit ein. Die Daten aus den verschiedenen Quellen bestätigten sich gegenseitig in ihrer Gültigkeit (*concurrent validity*).

Die Differenzierung von drei Bewertungsebenen in der ökonomischen Evaluation als Herzstück des HPM konnte durch diese Studie bestätigt und weiterentwickelt werden. Insbesondere wurden die beteiligten Parameter und die Zusammenhänge zwischen den drei Ebenen als verbundene Regulationskreise an einem Beispiel konkretisiert. Die Form der Visualisierung der Wirkungszusammenhänge in einem operationalen Strukturmodell erscheint wegweisend.

Der Prozess der ökonomischen Evaluation, wie es das HPM in sieben Arbeitsschritten vorsieht, konnte hier nicht untersucht werden, da es sich um eine reine Bewertungsevaluation handelte. Die Steuerungskomponente, also der Einfluss einer differenzierten Analyse und Planung vor der Durchführung einer Maßnahme, wurde nicht erfasst.

Es folgen zwei Evaluationsstudien, die zwar weniger aufwändig sind, die jedoch jede für sich eine besondere Herausforderung barg. Zunächst wird die Evaluation eines Personalentwicklungsprogramms eines *Call Centers* der öffentlichen Verwaltung dargestellt. Die besondere Herausforderung lag darin, dass keine konkrete Einzelmaßnahme, sondern ein Bündel verschiedener Interventionen im Verbund evaluiert werden sollte.

2.3 Programmevaluation: Personalentwicklungsmaßnahmen im Call Center

2.3.1 Hintergrund und Zielsetzung

Das Ziel der ersten (CISM, Kapitel 2.2) und der letzten (*E-learning*, Kapitel 2.4) Evaluationsstudie lag in der Analyse der Wirkung und Wirtschaftlichkeit konkreter Einzelmaßnahmen aus den Bereichen *Human Factors* und Training. Mit der in diesem Kapitel beschriebenen Evaluation wird dieser Ansatz erweitert. Untersucht werden soll die Wirkung mehrerer, aufeinander aufbauender Maßnahmen der Personalentwicklung für die Zielgruppe *Call Center* Agenten. Dabei soll im Vordergrund stehen, welchen Beitrag die einzelnen Maßnahmen für die beabsichtigten Wirkungen leisten und welche Bedeutung diese Wirkungen für die Leistungsfähigkeit des ausgewählten Organisationsbereichs haben. Aus dem HPM sollten damit die Arbeitsschritte „Funktionsanalyse“ und „Wirkungsanalyse“ umgesetzt und konkretisiert werden.

Untersucht wurden verschiedene Maßnahmen der Qualifizierung und Betreuung in einem *Call Center* der öffentlichen Verwaltung. Darüber hinaus wurden im Projekt F-2145 weitere empirische Evaluationen von Maßnahmen in *Call Centern* des öffentlichen und gewerblichen Bereichs durchgeführt (Pennig & Vogt, 2006).

Hintergrund für die Evaluation von Maßnahmen in *Call Centern* der öffentlichen Verwaltung war eine modular aufgebaute Praxishilfe, die im Rahmen des Projektes „VERT-iCall – Service Center in der öffentlichen Verwaltung gesund und effizient gestalten“ erarbeitet wurde. In dieser Praxishilfe werden Gestaltungshinweise und Instrumente für die Gestaltung der Arbeitsbedingungen, der Qualifizierung und der Führung beschrieben. Die Praxishilfe wird den *Call Centern* der öffentlichen Verwaltung durch die BAuA zur Verfügung gestellt. Das Projekt soll insbesondere die Nachhaltigkeit der Umsetzung ausgewählter Maßnahmen aus dieser Praxishilfe prüfen, um daraus eine Evaluationsroutine für vergleichbare Maßnahmen aus dem Bereich der Gesundheitsförderung zu entwickeln.

Im Rahmen der Auftrags- und Zielklärung mit den Führungskräften des *Call Centers* wurde deutlich, dass diese ein hohes Interesse daran hatten, die Wirksamkeit eines Bündels von Maßnahmen, die die Agenten in ihrer Arbeiten von der Einarbeitung bis zur fachlichen Supervision unterstützen sollen, zu untersuchen. Daraus entstand die Idee, eine alle qualifizierenden Maßnahmen integrierende Evaluation (Qualifizierungs-Programmevaluation) durchzuführen, die den Führungskräften außerdem ein Gesamtbild über die Bereiche Kompetenz und Leistung gibt.

2.3.2 Datenmodell für die Programmevaluation

Die Konzeption der programmbezogenen Evaluationsstudie sollte wie bereits die anderen Evaluationen auf dem *Human Resources Performance* Modell (HPM, Kapitel 1.4) basieren. Es wurde jedoch nicht alle sieben Schritte des HPM vollzogen, sondern lediglich die Funktions- und Wirkungsanalyse (Abbildung 1). Bei der Programmevaluation ergaben sich außer der reduzierten HPM-Vorgehensweise noch folgende Besonderheiten:

- (1) Ein Qualifikationsprogramm fördert vielfältige Kompetenzen (Wissen, Fertigkeiten, Fähigkeiten⁸) der Mitarbeiter. Die Programmevaluation erfordert daher umfassende Kompetenzmessungen.
- (2) Ein umfassendes Personalentwicklungsprogramm (PE-Programm), das nicht nur einzelne Fertigkeiten oder Verhaltenskompetenzen verbessern, sondern Mitarbeiter für die Arbeit „fit machen“ will, wird an dem Anspruch zu messen sein, dass die gesamte Tätigkeitsausübung und Leistung der Mitarbeiter gesteigert wird. Die Programmevaluation erfordert daher auch umfassende Leistungsmessungen.
- (3) Ein PE-Programm umfasst verschiedene Bausteine. Eine Evaluation sollte die spezifischen Effekte jedes einzelnen Bausteins aufzeigen. Die Programmevaluation erfordert demnach ein differenziertes Wirkungsmodell und eine Pfadanalyse.
- (4) Kompetenzen sind nur ein Faktor, der die Qualität der Tätigkeitsausübung und Leistung bestimmt; sie sind eine notwendige, aber keine hinreichende Bedingung für Erfolg. Um die Effekte des Programms gegen die Wirkungen anderer Faktoren abzugrenzen, sind auch die weiteren Einflussgrößen im Leistungssystem zu erfassen. Zu diesem System zählen alle direkten und indirekten Arbeitsbedingungen, wie Technik, Entgelte, Führung, Ausstattung, Information etc. Nur wenn diese Kontextfaktoren in der Evaluation erfasst werden, lassen sich Aussagen dazu treffen, ob das Personalentwicklungsprogramm seinen Beitrag zu optimaler Leistung beigetragen hat und welche anderen Faktoren, die Leistung darüber hinaus fördern oder begrenzen.
- (5) Die Bewertung des Leistungssystems und der systemfördernden Maßnahmen oder Programme erfordern einen Blick über die Kontextfaktoren hinaus. Diese Bewertung muss sich letztlich an den übergreifenden (ökonomischen und/oder strategischen) Kriterien der Organisation orientieren. Diese stellen den Steueringkontext des PE-Programms dar.

Die Struktur der PE-Programmevaluation zeigt das folgende Datenmodell (s. Abbildung 35).

⁸ Eine Fertigkeit ist ein auf der Grundlage von Fähigkeiten durch Lernen, Übung und Erfahrung ausgebildetes Können. Während Fähigkeiten häufig als angeborene Personeneigenschaften verstanden werden, sind Fertigkeiten konkrete, inhaltlich abgrenzbare und oft automatisierte Handlungen.

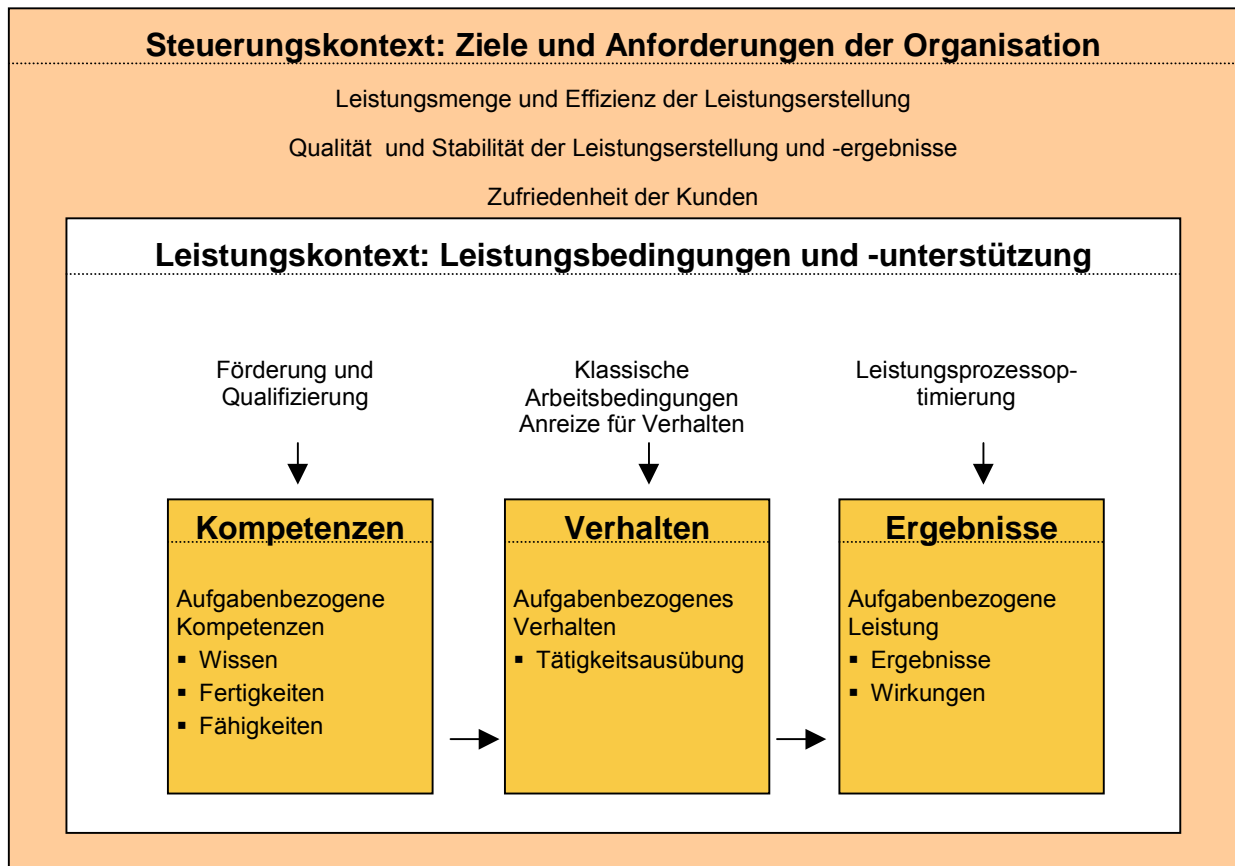


Abbildung 35: Datenmodell für die Programmevaluation

2.3.3 Planung und Konzeption der Programmevaluation

Das gesamte PE-Programm für Agenten in diesem *Call Center* umfasste fünf verschiedene, z.T. aufeinander aufbauende Maßnahmen:

1. Einarbeitung zu Beginn der Tätigkeit
2. Einweisung in Fachanwendungen (Fachanweisungen)
3. Kommunikationsschulung
4. Kurzanweisung zu aktuellen Auskunftsthemen
5. Teamentwicklungsmaßnahme

Die umfassende Programmevaluation sollte allerdings nur für die ersten drei Maßnahmen durchgeführt werden, da die Kurzanweisungen und die Teamentwicklungsmaßnahme nach Auskunft der Führungskräfte nur einen geringen Stellenwert für die Gesamtqualifizierung und Leistungsfähigkeit der Mitarbeiter haben. Allerdings sollte der Tatsache Rechnung getragen werden, dass sich die Agenten nach der Einarbeitung durch das „*learning-by-doing*“ weiter qualifizieren. Daraus ergab sich folgendes Pfadmodell zur Programmevaluation im untersuchten *Call Center* (s. Abbildung 36):

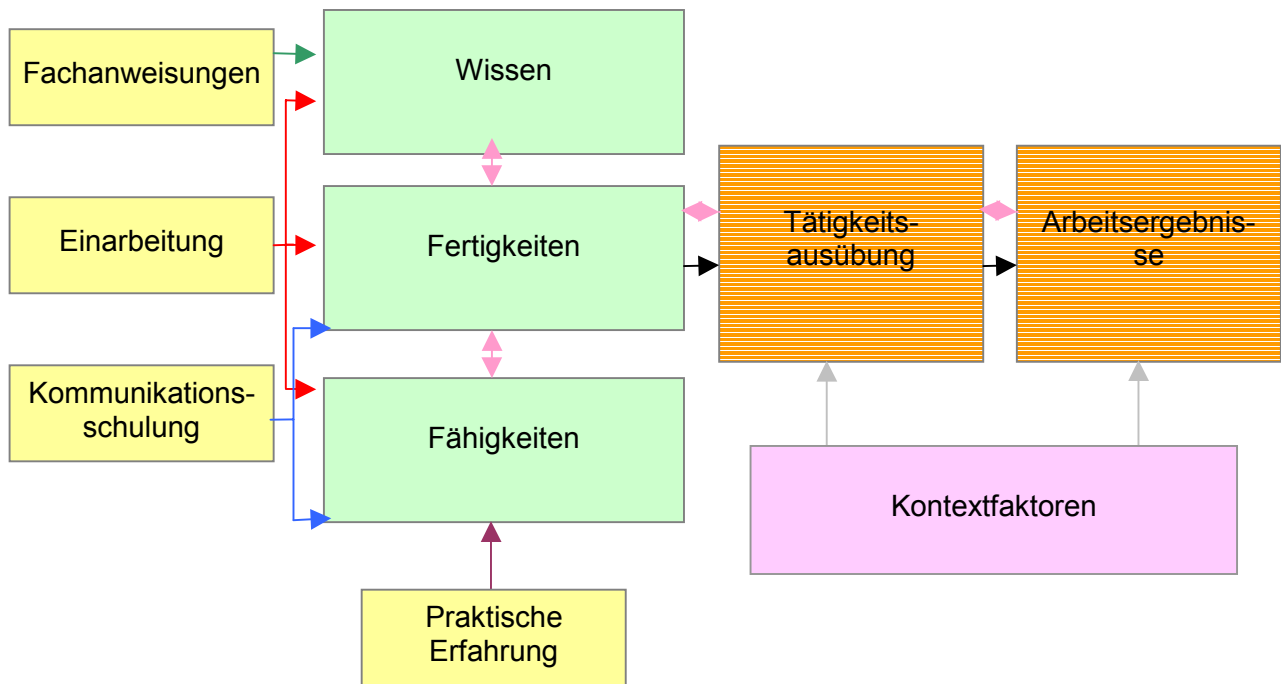


Abbildung 36: Pfadmodell zur Programmevaluation im untersuchten *Call Center*

Die farbigen Pfeile markieren die von den Führungskräften erwarteten und von den Autoren untersuchten Einflüsse der drei untersuchten Qualifizierungsmodule auf die verschiedenen Kompetenzbereiche. Dabei wurde davon ausgegangen, dass Fachanweisungen (grüner Pfeil) bei den *Call Center* Agenten überwiegend Wissen generieren, während die Kommunikationsschulung (blaue Pfeile) die Entwicklung von Fertigkeiten und Fähigkeiten unterstützt. Die Einarbeitung schließlich (rote Pfeile) sollte alle drei Bereiche fördern.

Die folgenden Fragestellungen sollten die Leitlinien zur Überprüfung der Wirksamkeit und Bedeutung des PE-Programms sein:

1. Welche Anforderungen sind besonders bedeutsam für die Beurteilung der Arbeit des *Call Center* Agenten (Inhalte der grünen und orangenen Felder)?
2. Entsprechen die durch die verschiedenen Maßnahmen entwickelten Kompetenzen (grüne Felder) den Anforderungen?
3. Welche Merkmale des Anforderungsprofils von Wissen, Fertigkeiten und Fähigkeiten werden durch die einzelnen Maßnahmen abgedeckt (farbige Pfeile)?
4. Wie effektiv werden durch die Maßnahmen die beabsichtigten Entwicklungen gefördert (farbige Pfeile in Verbindung mit den grünen Feldern)?
5. Welchen Effekt haben die initiierten und unterstützten Kompetenzentwicklungen auf die Tätigkeitsausübung und die Arbeitsergebnisse (schwarze Pfeile)?

6. Welchen Einfluss haben die Kontextfaktoren auf die Tätigkeitsausübung und die Arbeitsergebnisse (graue Pfeile)?
7. Wie sehr können durch Maßnahmen der Kompetenzförderung Ergebniseffekte erwartet werden (korrelationsanalytische Überprüfung des Pfadmodells; rosa Pfeile)?

Für die Programmevaluation wurden zwei Fragebogen konzipiert, die sich zum einen an die Agenten im *Call Center* (Selbsteinschätzung) und zum anderen an die Führungskräfte (u.a. Fremdeinschätzung) des *Call Centers* richteten. Beide Fragebogen beinhalteten sechs Abschnitte, die die verschiedenen Faktoren des Datenmodells abdecken (Abbildung 36):

- Abschnitt 1. Selbst-/Fremdeinschätzung des Wissens der Agenten
- Abschnitt 2. Selbst-/Fremdeinschätzung der Fertigkeiten der Agenten
- Abschnitt 3. Selbst-/Fremdeinschätzung der Fähigkeiten der Agenten
- Abschnitt 4. Bewertung der Tätigkeitsausübung und Aufgabenwahrnehmung
- Abschnitt 5. Effektivität und Auswirkungen der Arbeit der Agenten
- Abschnitt 6. Kontextfaktoren der Arbeit der Agenten

Anhand der Planungsgespräche und Dokumente im *Call Center* wurden die Anforderungskriterien für jeden Abschnitt zusammengestellt, mit den Führungskräften abgestimmt und als Items für die Befragung in Aussageform formuliert. Wie in der CISM-Hauptstudie (Kapitel 2.2) sollte die Ausprägung der Anforderungsmerkmale durch die Agenten (Selbsteinschätzung) und durch die Führungskräfte (Fremdwahrnehmung der durchschnittlichen Ausprägung im *Call Center*) anhand einer Prozentskala vorgenommen werden. Bei den Kontextfaktoren wurde der Grad des Einflusses jedes Faktors für die Tätigkeitsausübung auf einer 5er-Skala erfragt. Unterschieden wurden hier starke und schwache positive Einflüsse (++ oder +), kein Einfluss (o) und negative Einflüsse (- oder --).

Da das Evaluationsdesign keinen *Pretest* und keine Kontrollgruppe aufwies, wurden die Mitarbeiter und Führungskräfte direkt nach dem Zusammenhang von Anforderungsmerkmalen und Programmmodulen gefragt, um eine systematische Zuschreibung von Effekten zu ermöglichen. Es war anzugeben, welche Module auf welche Qualifikationsbereiche, Tätigkeiten und Ergebnismerkmale einen wirksam fördernden Einfluss gehabt haben. Außerdem sollte vermerkt werden, welche Kriterien als sehr bedeutsam für eine erfolgreiche Arbeit im *Call Center* gesehen werden, um eine Bewertung der Effekte vor allem aus der Sicht der Führungskräfte zu ermöglichen (Steuerungskontext).

Insgesamt waren 44 Anforderungskriterien hinsichtlich der drei Merkmale

- Ausprägung der Kompetenz (0-100%)
 - Bedeutung (ja/nein)
 - wirksame Förderung durch das Modul (ja/nein)
- einzuschätzen.

Bei den Kontextfaktoren wurde nicht nur nach deren Ausprägung, sondern auch nach den Auswirkungen auf die Anforderungen im Bereich der Tätigkeitsausübung und auf die Ergebnisse gefragt, um auch hier eine Einschätzung der Bedeutung von Effekten im Gesamtkontext vornehmen zu können.

Die inhaltsgleiche Struktur beider Fragebogen ermöglichte einen fast vollständigen Abgleich von Selbst- und Fremdwahrnehmung in der Evaluation.

2.3.4 Untersuchte Stichprobe und Rücklauf

Das für das Vorhaben gewonnene *Call Center* in einer Stadtverwaltung umfasst vier Teams mit insgesamt 45 Mitarbeitern, die Auskünfte erteilen bzw. als Vermittlungsstelle dienen. Es dient damit der Entlastung der verschiedenen Fachabteilungen und soll den Service der Verwaltung verbessern. Gelegentlich wird das *Call Center* auch als *Outbound-Center* genutzt (z.B. bei Bürgerbefragungen). Insgesamt waren 20 Mitarbeiter bereit, den Fragebogen auszufüllen. Ihre Teilnahme wurde mit einem Kinogutschein belohnt. Um die Qualität der Auskünfte angesichts des sehr umfangreichen und komplexen Fragenkataloges zu erhöhen, wurde der Fragebogen in Anwesenheit eines Projektmitarbeiters ausgefüllt, der Verständnisfragen beantworten konnte und sicherstellte, dass alle Agenten den gesamten Bogen bearbeiteten. Durch diese Form der betreuten Befragung war die Antwortqualität sehr hoch. Es kam selten vor, dass einzelne Fragenkomplexe nicht vollständig ausgefüllt waren. Als kleines Dankeschön erhielt jeder Studienteilnehmer einen Kinogutschein.

Alle verantwortlichen Führungskräfte des *Call Centers* (Leiter des *Call Centers* und seine Teamleiter) waren bereit, an der schriftlichen Befragung teilzunehmen. Somit lagen sechs Führungsfragebogen vor.

2.3.5 Ergebnisse

Die Darstellung der Ergebnisse erfolgt anhand der im vorletzten Kapitel 2.3.3 gestellten sieben Fragen. Zunächst wird dargestellt, welche Anforderungen eine besondere Relevanz für die Arbeit des Agenten haben und ob die Maßnahmen im PE-Programm Kompetenzen für die relevanten Anforderungen entwickeln. Die Kompetenzen (Wissen, Fähigkeiten, Fertigkeiten) auf der Funktionsebene des HPM (Kapitel 1.4 und Abbildung 1) sind jedoch nur eine notwendige, jedoch keine hinreichende Bedingung in der Wertschöpfung. Daher wurden auch die auf der Funktionsebene angesiedelten, individuellen Tätigkeiten und Leistungsergebnisse berücksichtigt.

2.3.5.1 Welche Anforderungen sind besonders bedeutsam für die Beurteilung der Arbeit des *Call Center* Agenten?

In einer ersten Analyse wurde überprüft, wie die verschiedenen Anforderungen im Hinblick auf ihre Bedeutsamkeit für ein erfolgreiches Arbeiten als Agent in dem *Call Center* eingeschätzt wurden. Dazu sollten die Items aus allen fünf Anforderungsbereichen ausgewählt werden, die als hoch erfolgskritisch angesehen wurden. Die Abbildung 37 zeigt, wie viele Items (prozentual) aus den fünf Bereichen ausgewählt wurden.

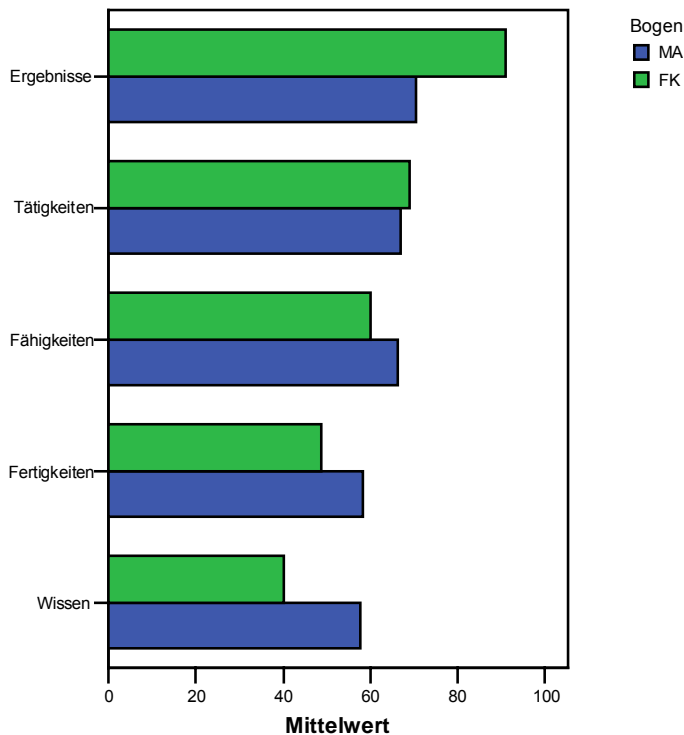


Abbildung 37: Prozentuale Anzahl von erfolgskritischen Items je Anforderungsbereich (MA: Mitarbeiter, FK: Führungskräfte)

Es zeigt sich ein sehr plausibles Ergebnis: Die Bedeutung der Items steigt im Pfadmodell von links nach rechts, also mit der Nähe zu den Ergebnisfaktoren der Arbeit. Das Wissen ist danach zwar eine wichtige Voraussetzung für erfolgreiches Arbeiten, wird aber weniger bedeutsam eingeschätzt als das Vorhandensein bestimmter Fähigkeiten oder die Qualität der Tätigkeitsausübung.

Dieses grundlegende Bedeutungsprofil wird von Führungskräften und Mitarbeitern gleichermaßen berichtet, wenn auch die Mitarbeiter ihre Kompetenzen (Wissen, Fertigkeiten, Fähigkeiten) wichtiger nehmen als die Führungskräfte.

Eine genauere Analyse auf Itemebene legte offen, welche konkreten Kompetenzen von Führungskräften und Mitarbeitern besonders (Nennungen von über 85%) hervorgehoben werden. Diese wurden in das Pfadmodell integriert (s. Abbildung 38):

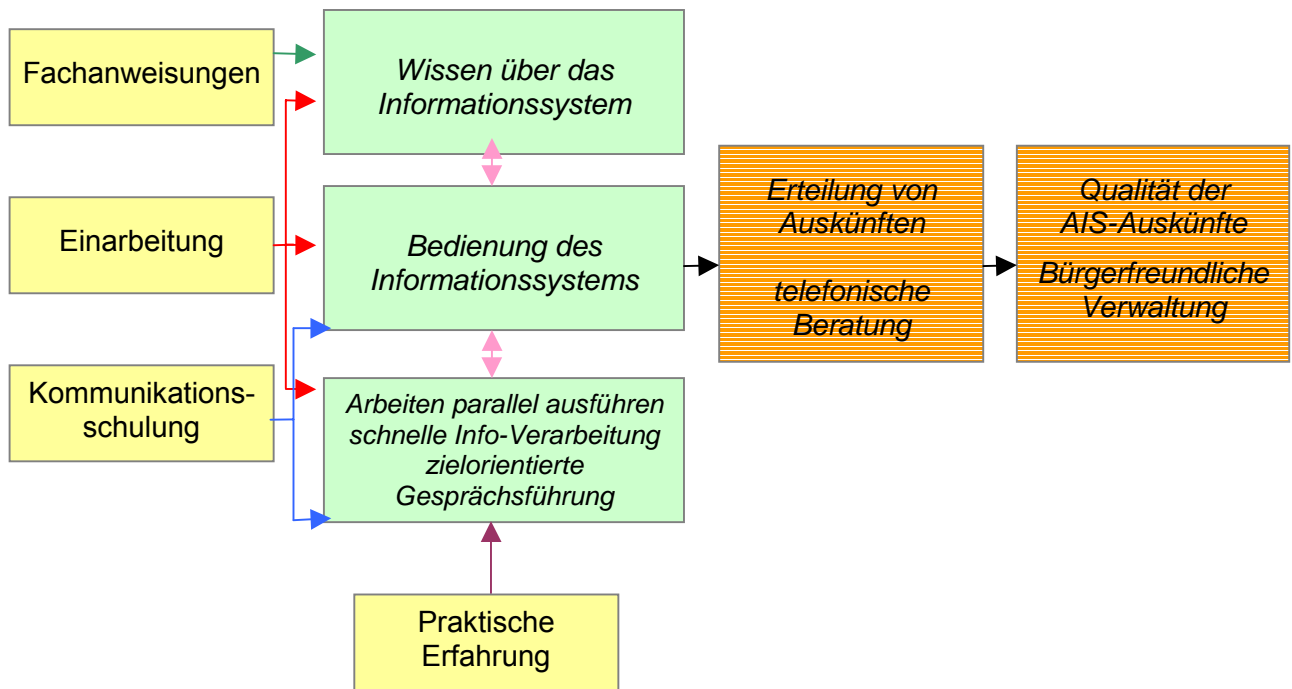


Abbildung 38: Erfolgskritische Anforderungen im *Call Center* dargestellt im Pfadmodell (AIS: Allgemeines Informationssystem)

Zentral für erfolgreiches Arbeiten im *Call Center* sind zunächst das Wissen über das allgemeine Informationssystem (AIS) und die Fertigkeiten, dieses bedienen zu können. In Verbindung mit den grundlegenden Fähigkeiten des Agenten parallel zu arbeiten, Informationen schnell zu verarbeiten und Gespräche zielorientiert zu führen, sind dies die wichtigsten Eignungsvoraussetzungen für die beiden Kerntätigkeiten: „Erteilung von Auskünften“ und „telefonische Beratung“. Das bedeutsamste Qualitätskriterium die Arbeit des Agenten betreffend ist die Qualität der Auskünfte mittels AIS. Dies zielt vor allem auf eine hohe Kundenzufriedenheit, die daran gemessen wird, ob der Bürger die Verwaltung bürgerfreundlich erlebt.

Diese neun Kriterien ergeben ein kompaktes Gesamtbild der zentralen Anforderungen an den *Call Center* Agenten in der öffentlichen Verwaltung.

2.3.5.2 Entsprechen die durch die verschiedenen Maßnahmen entwickelten Kompetenzen den Anforderungen?

Darauf aufbauend wurde in einem zweiten Schritt analysiert, welche Kompetenzen nach der Durchführung des Personalentwicklungsprogramms in welchem Umfang bei den Mitarbeitern vorhanden sind. Die Mittelwerte aller Kompetenzen in den drei Qualifikationsbereichen (Wissen, Fertigkeiten, Fähigkeiten) in der Selbst- und Fremdeinschätzung zeigt Abbildung 39 in Prozent. Eine Ausprägung von 75% wurde dabei als „hoch“ verankert; eine Ausprägung von 50% als „mittelmäßig“.

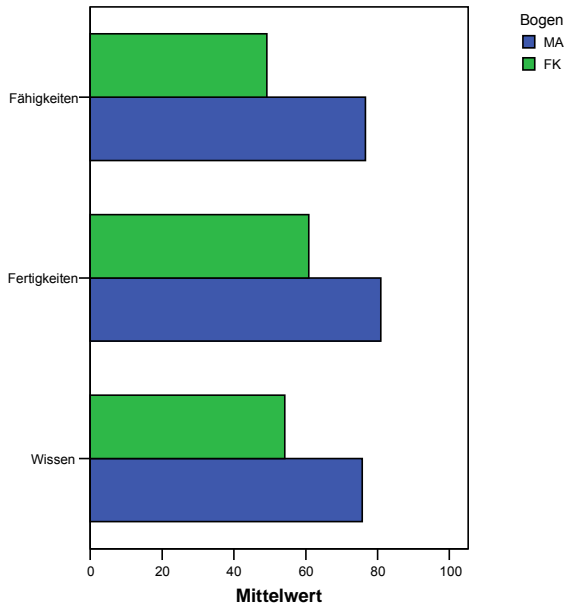


Abbildung 39: Kompetenzwerte für die drei Qualifikationsbereiche in der Selbst- und Fremdwahrnehmung (MA: Mitarbeiter, FK: Führungskräfte)

Insgesamt schätzen die Mitarbeiter ihre Qualifikationen durchweg sehr hoch ein und jeweils deutlich höher als die Führungskräfte, die insbesondere bei den grundlegenden Fähigkeiten am kritischsten sind.

Untersucht wurde daraufhin, welche konkreten Kompetenzen als besonders gut entwickelt und welche als besonders gering ausgeprägt erlebt werden.

Die gut ausgebildeten Kompetenzen bzw. Stärken der Mitarbeiter integriert Abbildung 40 in das Pfadmodell.

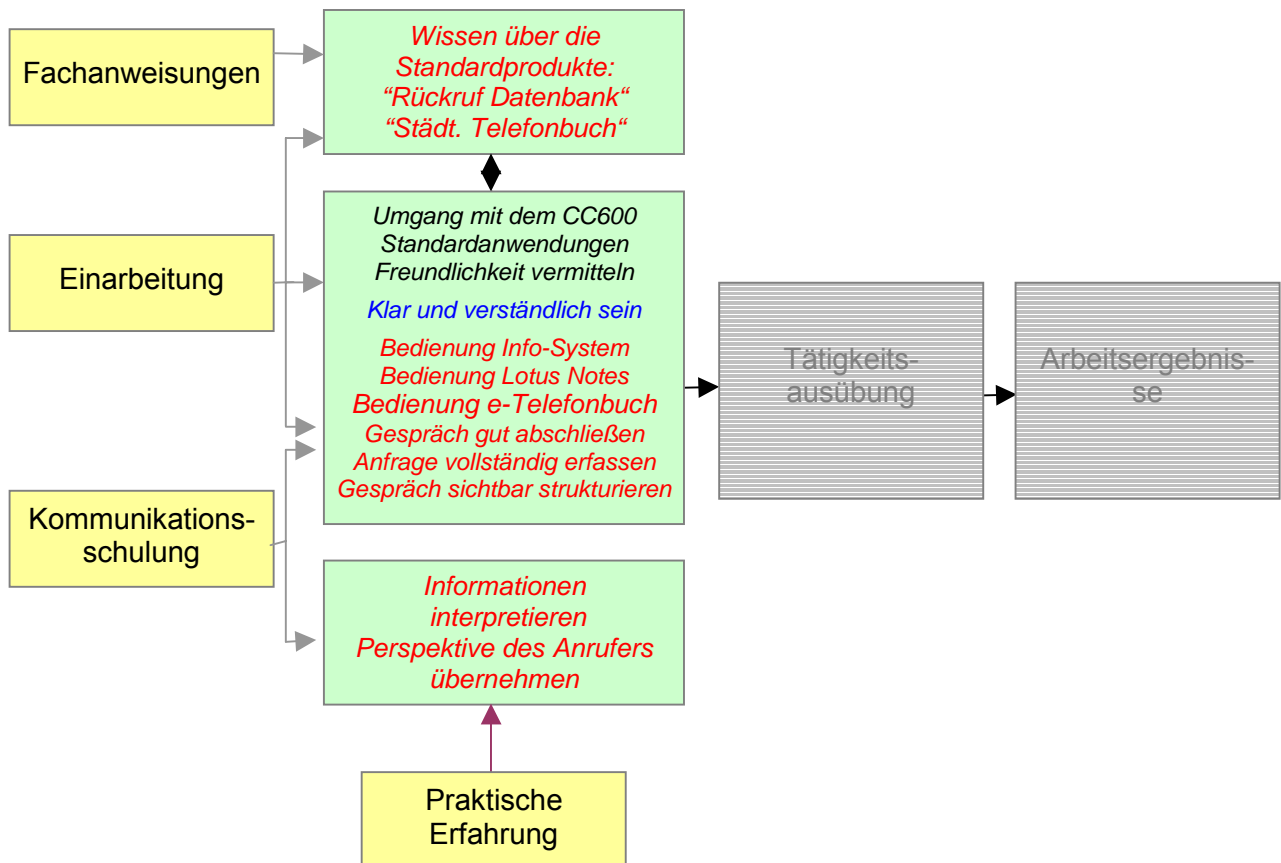


Abbildung 40: Hohe Kompetenzwerte im Call Center nach dem PE-Programm (blau: in Führungskräfte­sicht über 75%, rot: in Mitarbeitersicht über 80%, schwarz: in beiden Sichten hoch ausgeprägt)

In die Grafik wurden alle TOP-Werte der Führungskräfte (in blau; Werte über 75% Ausprägung) und der Mitarbeiter (in rot; Werte über 80% Ausprägung) aufgenommen (in schwarz: Merkmale, die von beiden Gruppen TOP-Werte bekamen).

Es zeigt sich, dass beide Gruppen die vorhandenen Fertigkeiten, insbesondere im Umgang mit den verschiedenen Unterstützungssystemen, als hoch einschätzen. Im Gegensatz zu den Führungskräften zählen die Mitarbeiter auch spezifische Wissensbereiche (Datenbanken) und grundlegende soziale Fähigkeiten (professioneller Umgang mit dem Anrufer) zu ihren Stärken.

Die Kompetenzen, die als nur mittelmäßig oder gering ausgeprägt gesehen werden, zeigt Abbildung 41 im Pfadmodell.

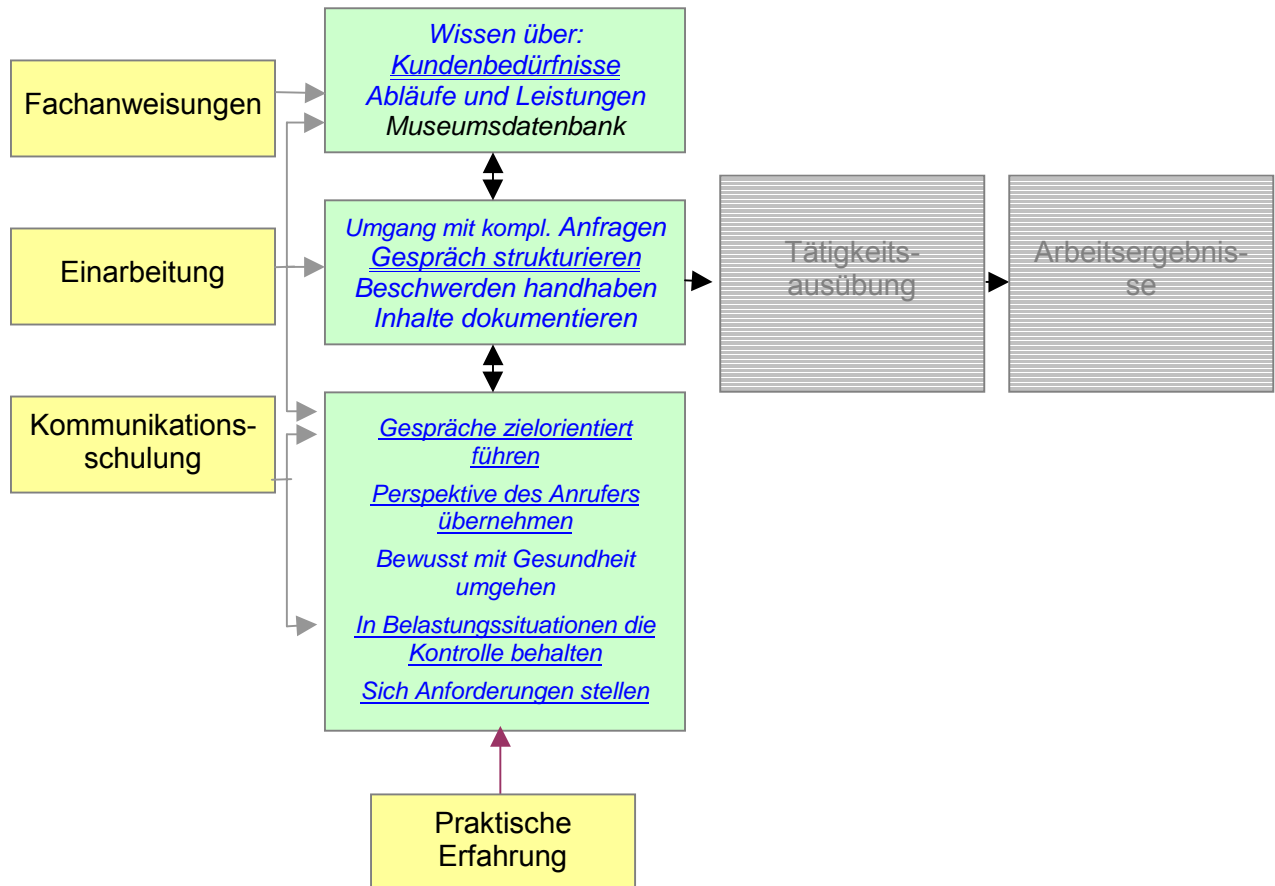


Abbildung 41: Niedrige Kompetenzwerte im Call Center nach dem PE-Programm (blau: in Führungskräfte-sicht unter 50%, schwarz: in beiden Sichten unter 60%, unterstrichen: erfolgskritisch in Führungskräfte-sicht)

Es gibt eine Reihe von Kompetenzen aus allen Bereichen, die von den Führungskräften trotz des PE-Programms als gering bis mittelmäßig ausgeprägt eingeschätzt werden (in blau; aufgeführt sind Werte < 50%). Auffällig häufig werden dabei grundlegende Fähigkeiten genannt. Die unterstrichenen Kompetenzen stellen die aus der Sicht der Führungskräfte erfolgskritischen Aspekte dar. Eine doppelte Unterstreichung zeigt an, dass alle Führungskräfte des Call Centers dieses Merkmal als sehr bedeutsam eingeschätzt haben.

Wie bereits in der Analyse der Kompetenzbereiche sehen die Mitarbeiter ihre eigenen Kompetenzen deutlich besser. Lediglich ein Kompetenzaspekt wird von ihnen als mittelmäßig eingestuft (Wert < 60%; in schwarz markiert, da er auch von den Führungskräften als gering eingestuft wurde).

Insgesamt werden also die Qualifikationen von Mitarbeitern und Führungskräften sehr unterschiedlich eingestuft. Es gibt eine Reihe von erfolgskritischen Merkmalen, die von den Führungskräften als „defizitär“ eingestuft werden.

2.3.5.3 Welche Merkmale des Anforderungsprofils werden durch die einzelnen Maßnahmen abgedeckt?

Im dritten Schritt der Evaluation des Personalentwicklungsprogramms sollte nun geklärt werden, welche der drei Module wirksam zur Förderung der Kompetenzen beigetragen hat (Zielbeitrag). Die Abbildung 42 zeigt, wie viel Prozent der Führungskräfte (in blau) bzw. der Mitarbeiter (in rot) dem jeweiligen Modul einen wirksamen Einfluss auf die Kompetenzbereiche zuschreiben.

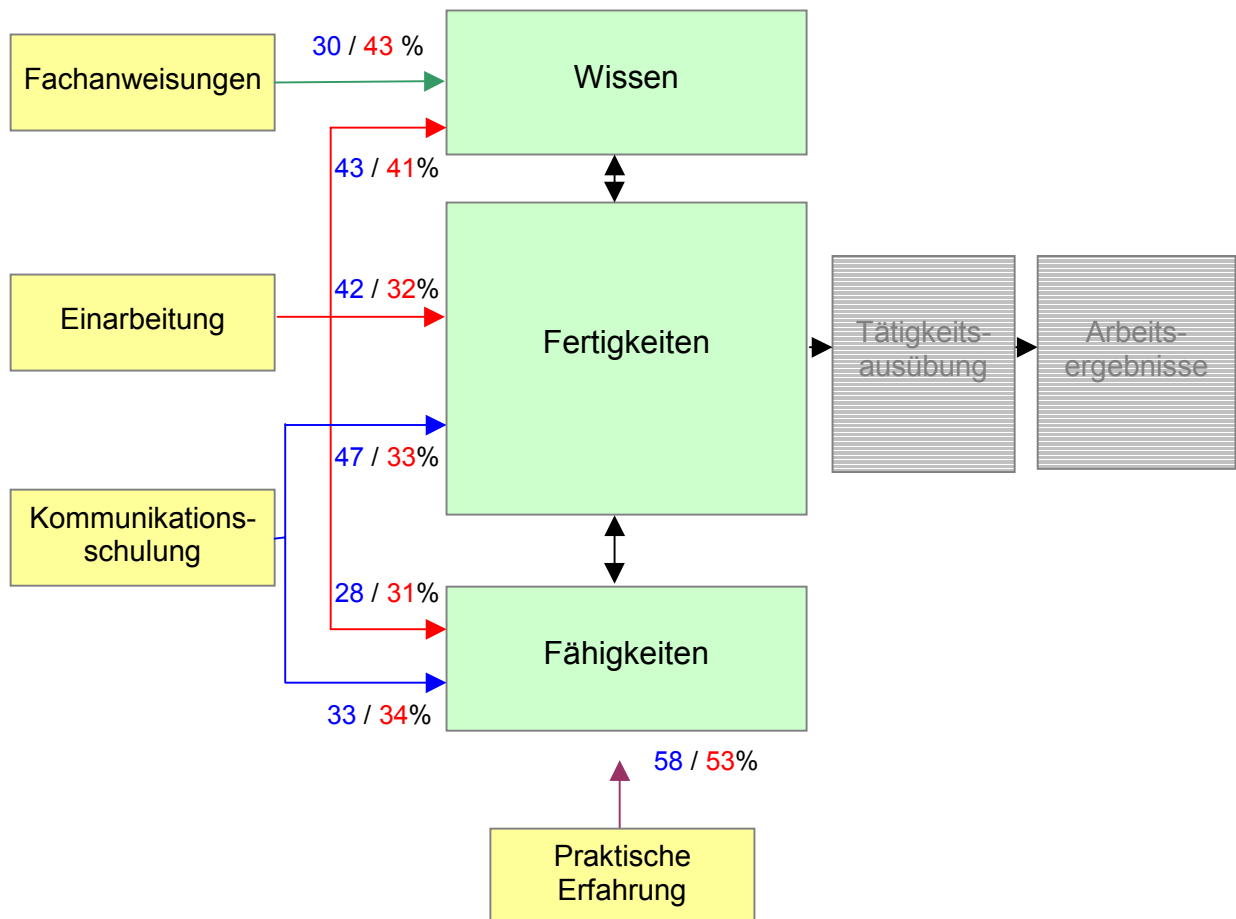


Abbildung 42: Zielbeiträge der Programmmodule zu den Kompetenzwerten im *Call Center* (blau: Mittelwert der Führungskräfte, rot: Mittelwert der Mitarbeiter)

Von beiden Gruppen wird der praktischen Erfahrung in der Arbeit im *Call Center* der höchste Kompetenzbeitrag (auf die Fähigkeiten) zugeschrieben. Es werden aber auch Differenzen im erlebten Zielbeitrag deutlich: Die Fachanweisung wird von den Mitarbeitern als (geringfügig) wichtiger erlebt als die Einarbeitung, während die Führungskräfte der Fachanweisung deutlich weniger Zielbeitrag für die Wissenskompetenz zuschreiben. Die Bedeutung der Einarbeitung und der Kommunikationsschulung für den Erwerb der Fertigkeiten wird von den Führungskräften höher eingeschätzt als von den Mitarbeitern. Die selbst wahrgenommene hohe Fertigkeitenkompetenz wird daher vermutlich von vielen Mitarbeitern eher der praktischen Erfahrung und damit der eigenen Lernleistung zugeschrieben.

Wichtig für die Evaluation ist des Weiteren, wie die Mitarbeiter den Zielbeitrag der einzelnen Module für ihre Qualifikation erleben. In Abbildung 43 sind die Zuschreibungen zu den einzelnen Kompetenzen aus Mitarbeitersicht im Pfadmodell dargestellt, die von mindestens 60% der Mitarbeiter angegeben werden. Die Farben signalisieren, welchen der drei Module der hohe Zielbeitrag jeweils zugeschrieben wird.

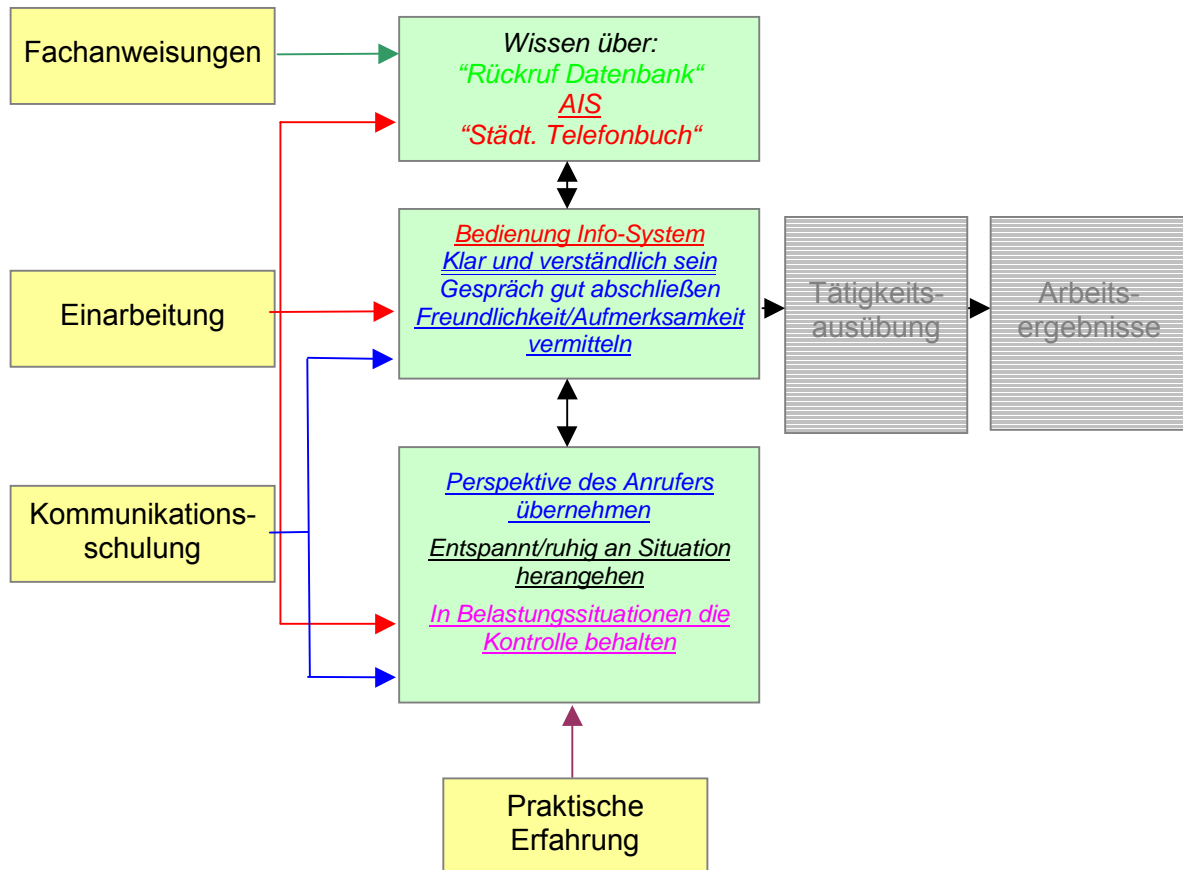


Abbildung 43: Anforderungsmerkmale mit hohen Zielbeiträgen der Programmmodule aus Mitarbeitersicht
(grün: Fachanweisung; rot: Einarbeitung; blau: Kommunikationsschulung;
schwarz: Kommunikationsschulung + praktische Erfahrung)

Auffällig ist, dass insbesondere die Mitarbeiter vor allem der Kommunikationsschulung hohe spezifische Zielbeiträge im Bereich der Fertigkeiten und Fähigkeiten zuordnen. Dies spricht dafür, dass diese Maßnahme als klar abgegrenzt und zielorientiert erlebt wird. Bei der Fachanweisung ist genau dies nicht der Fall. Ihr wurde zwar insgesamt ein hoher Zielbeitrag zum Wissen zugeschrieben, aber ein spezifisch hoher Zielbeitrag wird ihr nur in einem Anforderungsmerkmal zugeschrieben.

Abschließend zu diesem Punkt der Evaluation sollten die gravierendsten Abweichungen in den Zielbeiträgen untersucht werden: Bei welchen Kompetenzen sehen die Mitarbeiter geringe Zielbeiträge, obwohl die Führungskräfte den Modulen hohe Zielbeiträge (> 80%) zuschreiben? Große Abweichungen signalisieren, dass die Module vermutlich nicht die erwartete Kompetenzförderung erbringen. Die Anforderungsmerkmale mit großen Abweichungen zeigt Abbildung 44.

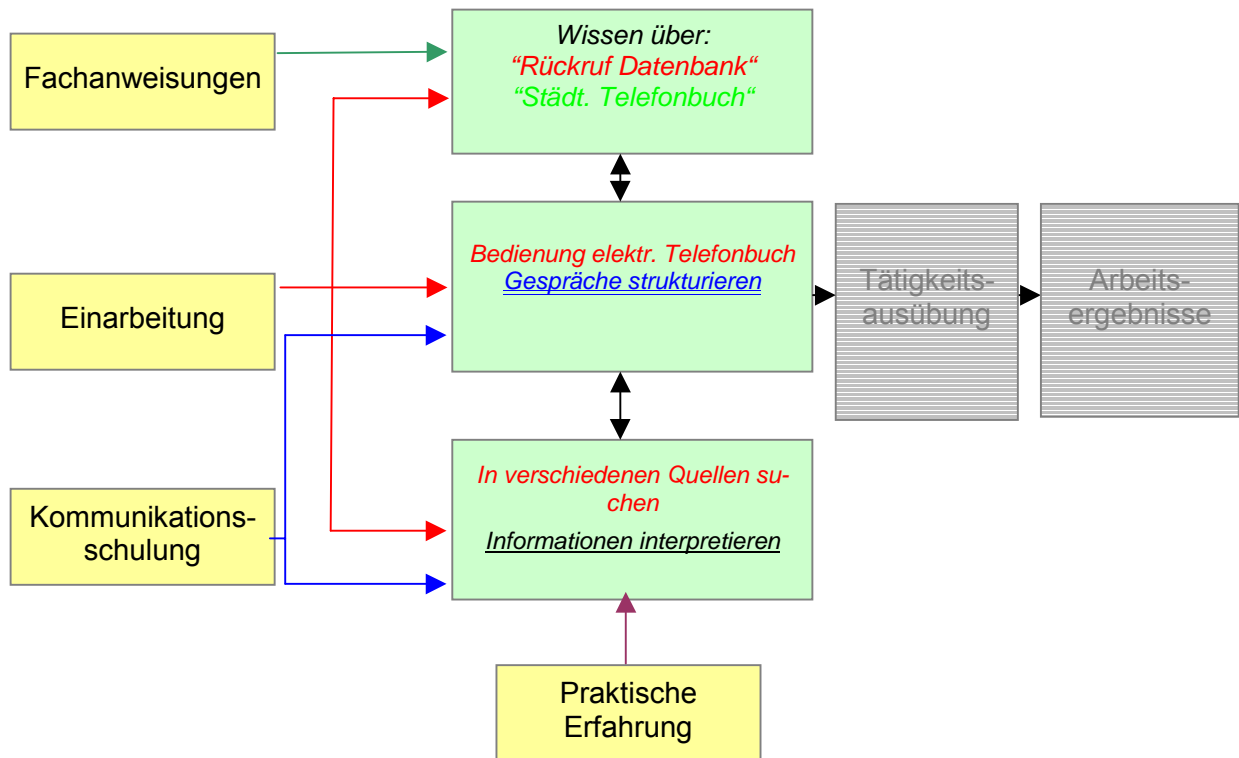


Abbildung 44: Kompetenzen, deren Realisierung durch das PE-Programm von Mitarbeitern und Führungskräften mit über 30% Differenz beurteilt wird
(grün: Fachanweisung; rot: Einarbeitung; blau: Kommunikationsschulung; schwarz: Kommunikationsschulung + praktische Erfahrung)

Es gibt zwei als erfolgskritisch eingestufte Merkmale, die durch das PE-Programm aus Sicht der Mitarbeiter weniger gefördert werden, als aus Sicht der Führungskräfte:

1. Die Kommunikationsschulung vermittelt nicht in dem Maße die Kompetenz, Gespräche strukturiert zu führen, wie es die Führungskräfte erwarten.
2. Die Fähigkeit, mündliche und schriftliche Informationen zu interpretieren, wird weder durch die Kommunikationsschulung noch durch die praktische Arbeit so vermittelt, wie es die Führungskräfte wahrnehmen.

Beide Punkte sind konkrete Hinweise zur Optimierung der Qualifikationsangebote.

2.3.5.4 Wie effektiv werden durch die Maßnahmen die beabsichtigten Entwicklungen gefördert?

Die bisherige Analyse hat gezeigt, dass insbesondere aus der Sicht der Führungskräfte bei einer Reihe von erfolgskritischen Anforderungskriterien die Kompetenzen eher gering ausgeprägt sind und daher ein Entwicklungsbedarf vorliegt. Die folgende Analyse soll aufzeigen, ob und welche Module des Entwicklungsprogramms für die Entwicklung dieser z.Zt. defizitären Kompetenzen eigentlich eingeplant sind.

Zu diesem Zweck wurden die Daten der Kompetenzwerte mit den Daten zum Zielbeitrag kombiniert. Das Ergebnis zeigt Abbildung 45.

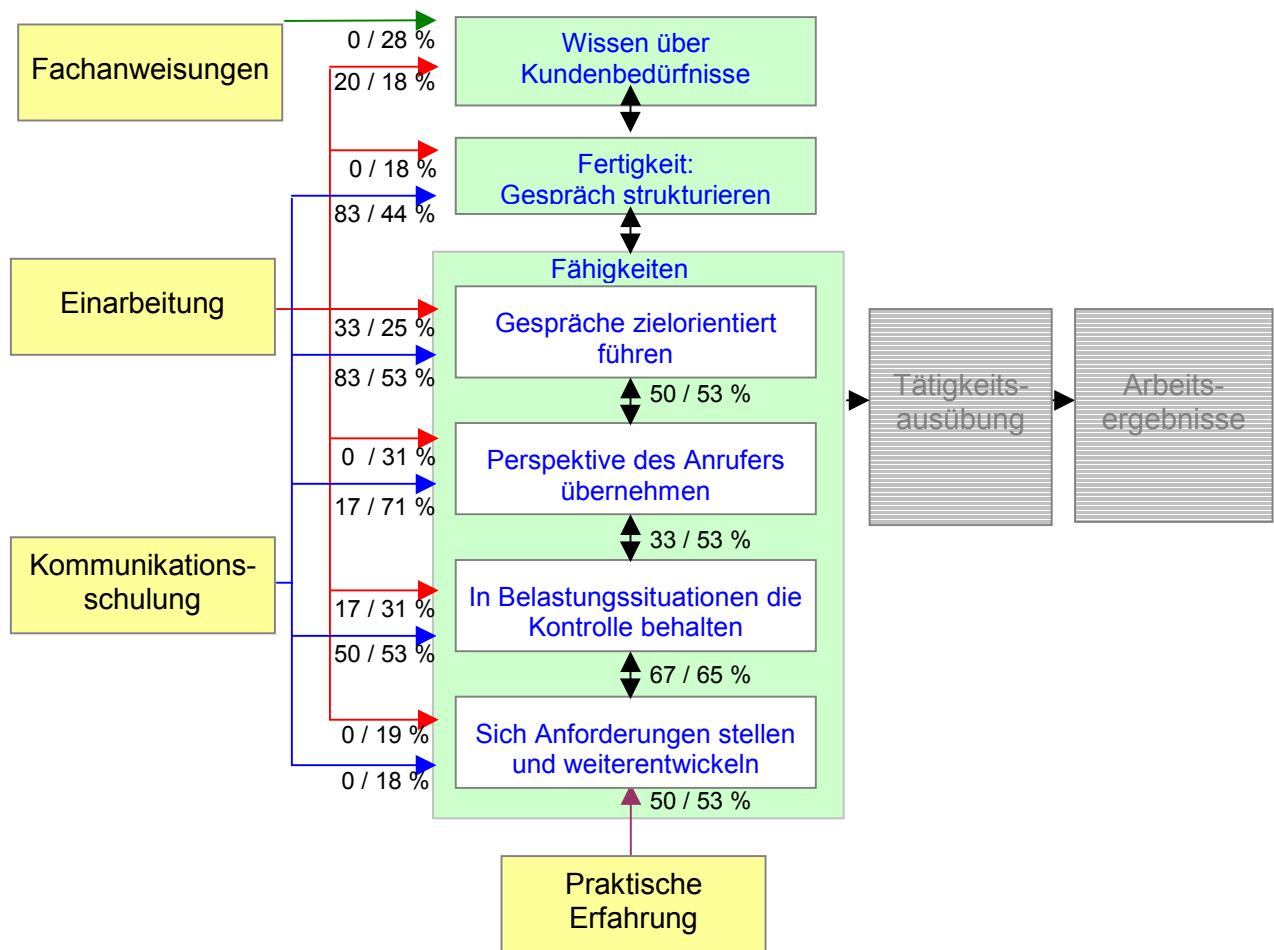


Abbildung 45: Zielbeiträge des Entwicklungsprogramms (Werte links der Schrägstriche) bei niedrigen Kompetenzwerten (Werte rechts der Schrägstriche) in erfolgskritischen Anforderungsmerkmalen aus Führungssicht

Es wurden für die Analyse die sechs Items ausgewählt, die aus Sicht der Führungskräfte als sehr bedeutsam erlebt werden und gleichzeitig niedrig ausgeprägt sind. Die Zielbeiträge aus der Wahrnehmung der Führungskräfte (jeweils erster Wert) und der Mitarbeiter (jeweils zweiter Wert) ist für jede erfragte Kombination von Maßnahme und Anforderungsmerkmal dargestellt.

Die Erkenntnisse dieser kombinierten Auswertung sind:

1. Die Führungskräfte erwarten mehr Wissen über die Bedürfnisse der Kunden, das *Call Center* bietet aber aus Sicht der Mitarbeiter hier wenig hilfreiche Unterstützung an. Das PE-Programm deckt dies kaum ab. Die bisherigen Qualifizierungsmaßnahmen sind vermutlich dazu eher ungeeignet.
2. Eine Kompetenz, die in engem Zusammenhang mit dem Wissen über Kundenbedürfnisse steht, ist die Perspektivenübernahme. Während die Führungskräfte diesbezüglich gleichzeitig ihren Anspruch und ihre kritische Einschätzung des Zielbeitrages aller Maßnahmen ausdrücken, erleben die Mitarbeiter die Kommunikationsschulung als hilfreich. Die Daten sprechen für die Notwendigkeit, die Bedeutung, die konkreten Erwartungen der Führungskräfte

und die Möglichkeiten einer systematischen Unterstützung zum Komplex „Kundenorientierung“ im Dialog und in der Führungsarbeit zu thematisieren.

3. Die als defizitär eingeschätzten Fertigkeiten und Fähigkeiten werden aus der Sicht der Führungskräfte zum Teil in hohem Maße durch die Kommunikationsschulung unterstützt, ohne aber ausreichende Effekte zu erzielen. Auch die Mitarbeiter sehen hier einen guten Beitrag der Kommunikationsschulung. Das PE-Programm sollte aufgrund dieser Erkenntnisse den Bereich der Verhaltensschulung in dem PE-Programm verstärken.
4. Es wird von den Führungskräften erwartet, dass ihre Mitarbeiter aktiver Herausforderungen suchen und initiativ lernen. Eine Anleitung zum Lernen bieten die Einarbeitung und die weiteren Module jedoch nicht. Anspruch und Wirklichkeit liegen hier entsprechend auseinander.

2.3.5.5 Welchen Effekt haben die initiierten und unterstützten Kompetenzentwicklungen auf die Tätigkeitsausübung und die Arbeitsergebnisse?

Letztlich wird das Personalentwicklungsprogramm daran gemessen, welche Effekte die Maßnahmen und damit die entwickelten Kompetenzen auf die Qualität der Tätigkeitsausübung und die Qualität der Arbeitsergebnisse haben.

Dazu sollte zunächst überprüft werden, wie die Qualität der Tätigkeitsausübung und der Arbeitsergebnisse nach dem PE-Programm von Führungskräften und Mitarbeitern beurteilt werden. Die Abbildung 46 gibt die Beurteilungswerte für diese beiden Ergebnisbereiche wieder.

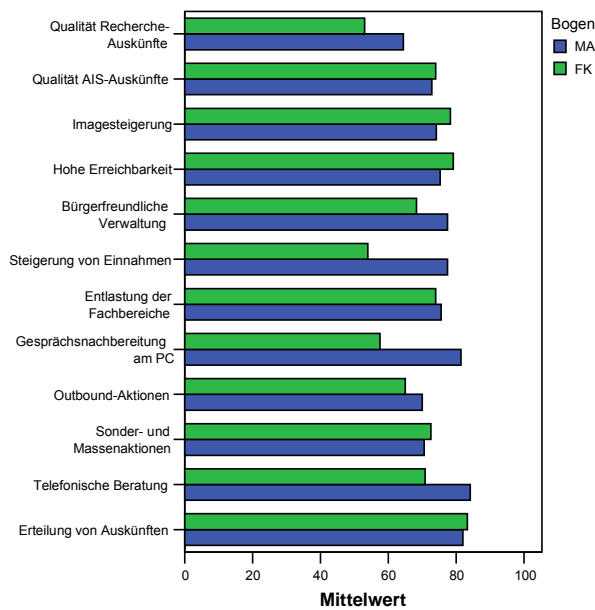


Abbildung 46: Qualität der Tätigkeitsausübung und der Ergebnisse aus Sicht der Führungskräfte (FK) und Mitarbeiter (MA)

Ähnlich wie bei den Kompetenzen äußern sich die Führungskräfte bei vielen Punkten kritischer als die Mitarbeiter. In beiden Gruppen wird jedoch die zentrale Tätigkeit „Erteilung von Auskünften“ im *Call Center* gleichermaßen positiv eingeschätzt (>80%). Kritischer sind die Führungskräfte vor allem bei den Tätigkeiten „Gespräche nachbereiten“, „Auskünfte nach ausführlicher Recherche geben“, „telefonisch beraten“, „*Outbound*-Aktionen durchführen“ sowie „Einnahmen durch weitere Kunden steigern“. Zusammenhänge zwischen den beiden problematischen Ergebnisbereichen „Recherche-Qualität“ und „Qualität der telefonischen Beratung“ und dem als defizitär erlebten Kompetenzkomplex „Kundenorientierung“ sind offensichtlich.

Um den Zusammenhang zwischen Ergebnisqualität und PE-Programm herzustellen, wird im Folgenden untersucht, welche Bedeutung die verschiedenen Programmmodule für eine erfolgreiche Tätigkeitsausübung und eine hohe Ergebnisqualität haben. Den von Mitarbeitern und Führungskräften hier zugeschriebenen Zielbeitrag zeigt Abbildung 47.

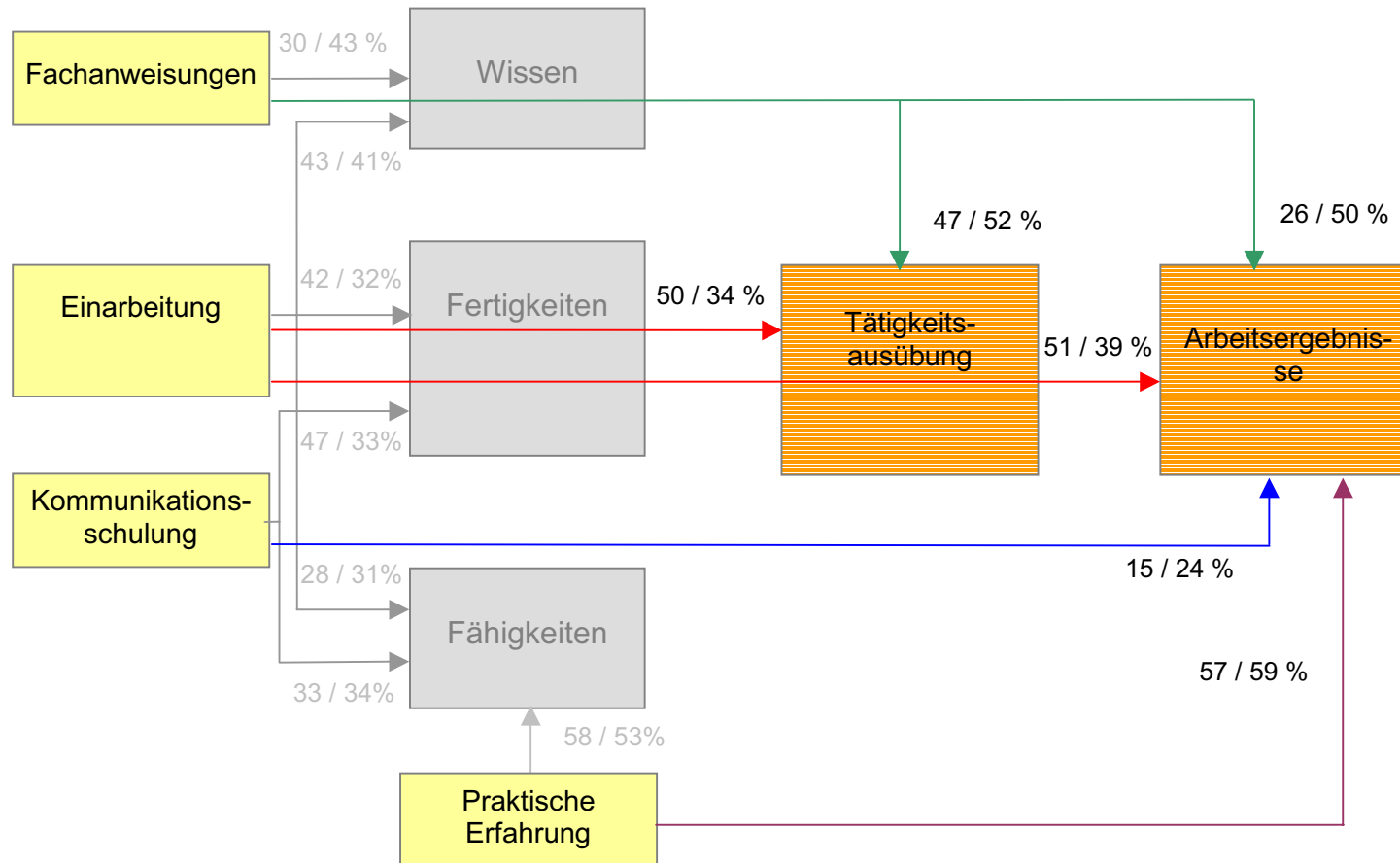


Abbildung 47: Zielbeiträge der Programmmodule zu Tätigkeits- und Ergebnisqualität aus Sicht der Führungskräfte (Werte links der Schrägstriche) und der Mitarbeiter (Werte rechts der Schrägstriche); die entsprechenden Werte der Kompetenzbereiche in hellgrau zeigen, dass die Befragten Zusammenhänge von Kompetenzen, Tätigkeits- und Ergebnisqualitäten realisieren

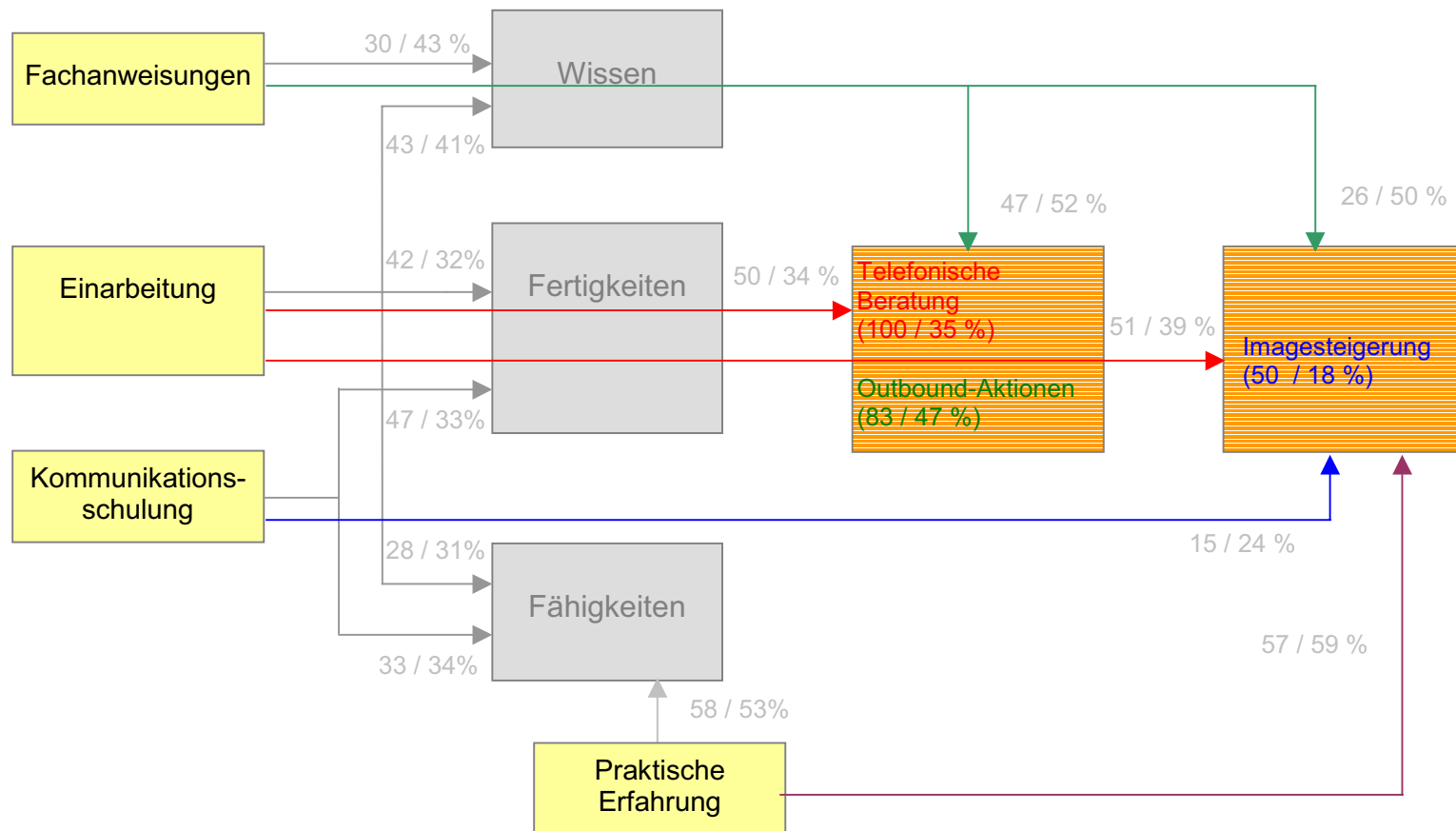


Abbildung 48: Tätigkeits- und Ergebnismerkmalen, deren Realisierung durch das PE-Programm aus Sicht der Führungskräfte (Werte links der Schrägstriche) und der Mitarbeiter (Werte rechts der Schrägstriche) über 30% abweicht

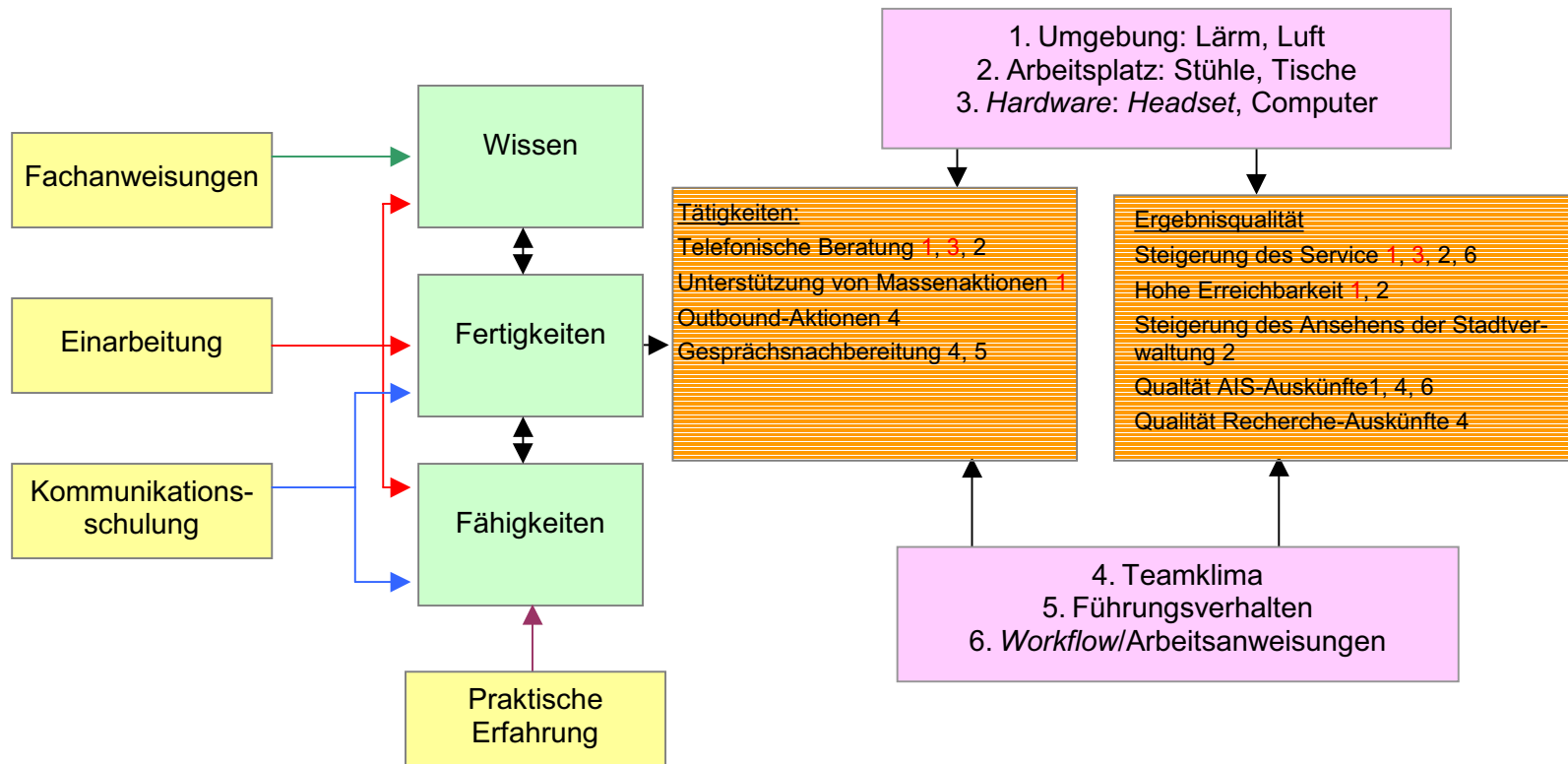


Abbildung 49: Einflüsse der Kontextfaktoren auf die Tätigkeitsausübung und die Ergebnisqualität
(Zahlen als Querverweis zu den Kontextfaktoren; rot = negativer Einfluss)

Die Daten vervollständigen die bisherigen Erkenntnisse:

1. Es bestätigt sich, dass der Kommunikationsschulung bisher (zu) wenig Bedeutung für die Ergebnisqualität der Arbeit in diesem *Call Center* zugestanden wird, obwohl das Potenzial gesehen wird, die zentralen erfolgskritischen Fertigkeiten und Fähigkeiten, die zum Teil bisher defizitär sind, entscheidend zu fördern.
2. Der Fachanweisung wird für die Tätigkeitsausübung und für die Qualität der Arbeit von den Mitarbeitern mehr Zielbeitrag zugeschrieben als der Einarbeitung, was die Führungskräfte genau umgekehrt sehen.
3. Für die Beurteilung des PE-Programms ist es ernüchternd, dass gerade der praktischen Erfahrung, also dem nicht angeleiteten Lernen, von beiden Seiten der größte Zielbeitrag für die Ergebnisqualität zugeschrieben wird.

Auch hier sollte eine Analyse auf Itemebene Aufschlüsse darüber liefern, bei welchen Anforderungsmerkmalen die Abweichungen zwischen der Führungs- und der Mitarbeitersicht besonders hoch sind. Die Merkmale mit Abweichungswerten über 30% zeigt Abbildung 48.

Es konnten drei große Abweichungen gefunden werden:

1. Die Einarbeitung trägt nicht in dem Maße zur Qualität der telefonischen Beratung bei, wie es die Führungskräfte erwarten.
2. Die Fachanweisung qualifiziert aus Sicht der Mitarbeiter nicht in dem Maße für *Outbound*-Aktionen, wie es sein sollte.
3. Die aus der Kommunikationsschulung resultierenden Verhaltenskompetenzen haben aus Mitarbeiterperspektive insbesondere nicht den positiven Effekt auf die *Imageverbesserung* der Stadtverwaltung wie es die Führungskräfte erwarten. Dies macht noch einmal deutlich, dass die Bedeutung der kommunikativen Fähigkeiten durch die Führungsarbeit verstärkt vermittelt werden sollte und sich dies stärker im PE-Programm niederschlagen müsste.

2.3.5.6 Welchen Einfluss haben die Kontextfaktoren auf die Qualität der Arbeit?

Zur Vervollständigung der Programmevaluation wurde auch nach dem Einfluss der Kontextfaktoren auf die Tätigkeitsausübung und die Ergebnisqualität gefragt. Die Ergebnisse zeigt Abbildung 49.

Insbesondere die Umgebung (Lärm und schlechte Luft) wird von den Mitarbeitern als störender Kontextfaktor genannt, der sich auf die telefonische Beratung, die Unterstützung von *Outbound*-Aktionen, die Steigerung des Services und die Erreichbarkeit auswirkt. Aber auch das *Headset* wird von einigen Mitarbeitern als hinderlich für gute Beratung und guten Service genannt („Verständigungsprobleme“). Dies sind wichtige Hinweise zur weiteren Optimierung der Kerntätigkeit der Agenten. Wenn z.B. im Verhaltensbereich weitere Unterstützungsmaßnahmen angeboten werden, sollte dies

auch durch eine bessere *Headset*-Technik begleitet sein, damit neu entwickelte Kompetenzen nicht untergehen.

Besonders positiv wirkt sich das Teamklima auf die Arbeit aus, z.B. auf die Qualität der Auskünfte. Vermutlich hilft hier die gegenseitige Unterstützung als Ergänzung zu den angebotenen Programmmodulen.

2.3.5.7 Wie sehr können durch Maßnahmen der Kompetenzförderung Ergebniseffekte erwartet werden?

Mit dem dargestellten Pfadmodell wurde die Annahmen verbunden, dass die Kompetenzförderung einen wichtigen Beitrag für die Verbesserung der Ergebnisqualität im *Call Center* leistet. Eine Korrelationsanalyse sollte abschließend aufzeigen, ob diese vermuteten Zusammenhänge auch durch die Daten gestützt werden. Eine Analyse der Korrelationen der Anforderungsbereiche ergab folgende Zusammenhänge (s. Abbildung 50).

Die Korrelationsanalyse zeigt auf, dass es z.T. sehr starke Zusammenhänge zwischen den Kompetenzen und den Ergebnismerkmalen in der erwarteten Konstellation gibt. Die grundlegenden Fähigkeiten weisen von den drei Kompetenzbereichen den stärksten Zusammenhang mit der Qualität der Tätigkeitsausübung und der Ergebnisqualität auf. Wissen und Fertigkeitenkompetenzen weisen dagegen so gut wie keinen Zusammenhang zur Ergebnisqualität auf. Die Korrelationen bestätigen das Pfadmodell und unterstützen die Empfehlung aus der bisherigen Analyse, im Bereich der Fähigkeiten systematischer und nachhaltiger anzusetzen.

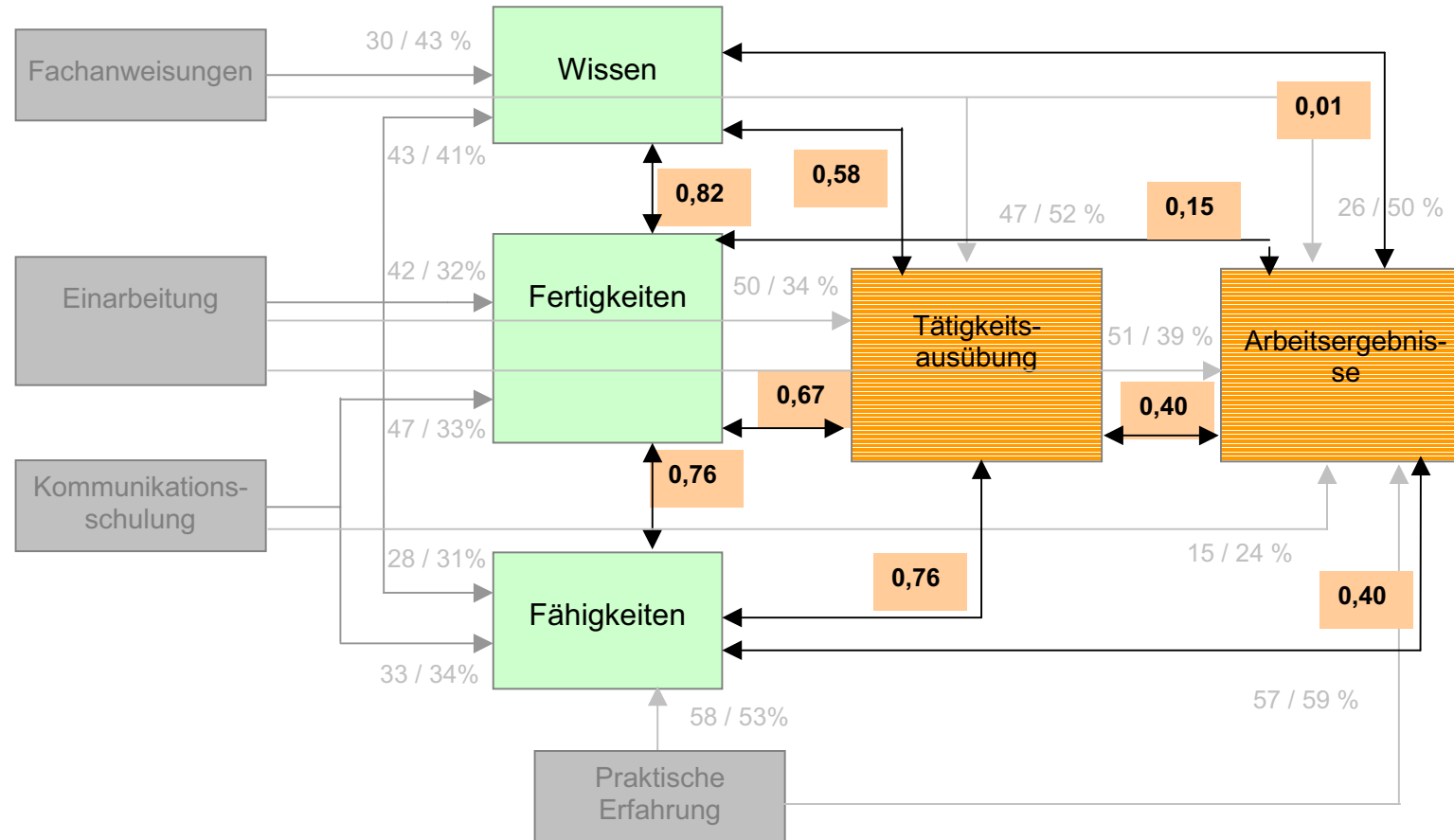


Abbildung 50: Korrelationen im Pfadmodell

2.3.6 Zusammenfassung und Resumée zur Programmevaluation

Mit dem zweiten Evaluationsprojekt sollte die Wirkung mehrerer, aufeinander aufbauender Maßnahmen der Personalentwicklung für die Zielgruppe „*Call Center* Agenten“ untersucht werden. Ziel war es, den Beitrag des Gesamtprogramms und der einzelnen Programmmodule für die Kompetenzentwicklung des einzelnen Mitarbeiters und seine Leistungsfähigkeit im Gesamtsystems „*Call Center*“ nachzuweisen. Die Evaluation zielte damit auf die Funktions- und Wirkungsanalyse aus dem HPM.

Um die hohe Komplexität des Evaluationsvorhabens handhabbar zu machen, wurde zunächst ein Datenmodell sowie ein Pfadmodell erarbeitet, die die Grundarchitektur der Programmevaluation in der Personalentwicklung, die zu berücksichtigenden Konstruktbereiche sowie deren Wechselwirkungen aufzeigen sollte und als Leitlinie für die Planung, Konzeption und Auswertung des Vorhabens fungierte. Dieses Vorgehen hat sich für die Praxis bewährt, weil es die Zusammenhänge zwischen den einzelnen Maßnahmen, veränderten Kompetenzen, Tätigkeitsausübungen und Ergebnissen für das Gesamtsystems „*Call Center*“ plausibel darstellen konnte. Daraus lässt sich ableiten, welche Maßnahme in dem Gesamtpaket besonders stark mit erwünschten Veränderungen korreliert und welche hinter den Erwartungen zurückbleibt. Dies ermöglicht unmittelbar die Optimierung des Programms und seiner Anwendung (Führung). Aus wissenschaftlicher Sicht hat das korrelative Vorgehen zwar den Nachteil, dass Ursache-Wirkungs-Beziehungen zwischen den Programmteilen und den Wirkungen nicht nachgewiesen werden konnten. Die mangelnde Kausalitätsanalyse ist für Praktiker jedoch wenig relevant, weil sie vorwiegend daran interessiert sind, das Richtige zu tun und nicht unbedingt wissen müssen, warum es richtig ist.

Die Evaluation wurde durch eine schriftliche Befragung der Mitarbeiter und der Führungskräfte realisiert. Dank einer sehr hohen Rücklaufquote von 44% bei den Mitarbeitern (20 Fragebogen) und 100% bei den Führungskräften (6 Fragebogen) und der hohen Datenqualität stand eine solide Stichprobe mit einem aussagekräftigen Datenmaterial für die Auswertung zur Verfügung.

Die Auswertung wurde anhand von sieben aufeinander aufbauenden Fragestellungen durchgeführt. Die Ergebnisse zeigten sowohl den Nutzen des Personalentwicklungsprogramms als auch den Optimierungsbedarf auf.

Das Personalentwicklungsprogramm vermittelt vor allem durch eine systematische Einarbeitung den Umgang mit den Standardanwendungen und dem Informationssystem. Dies sind Kompetenzen, die von beiden Seiten durchweg als sehr bedeutsam für das erfolgreiche Arbeiten im *Call Center* eingestuft werden. Dass hier allerdings nur die Grundkompetenzen vermittelt werden, zeigte sich darin, dass die Mitarbeiter den regelmäßigen Fachanweisungen mehr Beitrag zu ihrer Förderung im Bereich des Wissens und mehr Relevanz für eine erfolgreiche Tätigkeitsausübung zuschreiben als der Einarbeitung.

Aus der Sicht der Führungskräfte sind viele Qualifikationen nach dem Personalentwicklungsprogramm im Durchschnitt nur mittelmäßig vorhanden. Dazu zählen auch eine Reihe von sehr bedeutsamen Wissensbereichen, Fertigkeiten und vor allem Fä-

higkeiten, die für den Umgang mit dem Kunden und eine gezielte Beratung Voraussetzung sind.

Die Ergebnisse offenbaren, dass der Anspruch der „Kundenorientierung“ mit dem bisherigen PE-Programm nicht ausreichend eingelöst wird. Die Führungskräfte erwarten, dass die Kommunikationsschulung dies abdeckt, aus der Sicht der Mitarbeiter zeigen sich aber diesbezüglich nur sehr begrenzte Lernerfolge, die diesem umfassenden Anspruch nicht gerecht werden. Sie wirken sich daher auch nicht entscheidend auf die Ergebnisqualität aus. Stattdessen wird besonders der Lernbeitrag des „*learning-by-doing*“ als notwendig und wichtig erlebt. Letztlich wird die „strategische Bedeutung“ der Kommunikationsförderung für die Ergebnisqualität von den Führungskräften als eher gering eingeschätzt, was eine mögliche Ursache für die nicht ausreichende Verhaltensunterstützung sein könnte. Hinzu kommt, dass ein tieferes Verständnis der Kundenperspektive durch die bisherigen Maßnahmen kaum vermittelt wird. Die Korrelationsanalyse bestätigt die These, dass die Ausprägung der grundlegenden Fähigkeiten der Mitarbeiter eng mit der Ergebnisqualität zusammenhängt. Insgesamt ist es für das PE-Programm ernüchternd, dass dem Lernen durch Erfahrung der höchste Beitrag für die Ergebnisqualität zugeschrieben wird.

Die Kontextfaktoren der Arbeit werden von den Mitarbeitern weitestgehend als förderlich erlebt. Lärm, schlechte Luft und ein unzureichendes *Headset* erschweren jedoch aus Sicht vieler Mitarbeiter die Verständigung mit dem Anrufer und eine gute Servicequalität.

Mit dieser Form der Programmevaluation konnten somit wertvolle Erkenntnisse zur Wirksamkeit und Wirkungsweise des PE-Programms und der einzelnen Module gewonnen werden. Die komplexe Wirkungskette konnte mit einer innovativen Methodik auch ohne ein wissenschaftlich solides Design (*Pretest* oder Kontrollgruppe) abgebildet und nachgewiesen werden. Insofern zeigt diese Studie einen methodischen Ansatz auf, der die Programmevaluation für die Praxis leichter umsetzbar macht.

Gleichwohl basiert diese Methodik auf einem komplexen Datenmodell, das ein profundes Verständnis der möglichen Wirkungszusammenhänge erfordert und an die Befragten hohe Anforderungen stellt. Die Qualität der ausgefüllten Fragebogen spricht allerdings dafür, dass die Befragten mit der Komplexität der geforderten Auskünfte umgehen konnten.

Insgesamt konnten mit dieser Untersuchung die Arbeitsschritte „Funktionsanalyse“ und „Wirkungsanalyse“ aus dem HPM (Kapitel 1.4) konkretisiert werden. Das erarbeitete Pfadmodell kann als grundlegendes Evaluationsschema für zukünftige programmevaluatorische Untersuchungen dienen. Es wäre darüber hinaus sehr wertvoll, wenn die aufgezeigte Methodik zukünftig im Rahmen eines quasi-experimentellen Designs mit Kontrollgruppe oder *Pretest* validiert werden könnte, damit auch die den dargestellten Zusammenhängen zugrunde liegenden Kausalbeziehungen nachgewiesen werden können. Trotz fehlender Kausalaussagen demonstriert die Programmevaluation, dass das HPM nicht nur auf einzelne HF/HR/T-Maßnahmen, sondern auch auf Maßnahmenbündel angewendet werden kann und zu wertvollen Ergebnissen führt.

Das folgende Kapitel führt uns zur dritten und letzten empirischen Studie dieses Berichts: einer Kurzevaluation von *E-learning* Modulen. Ziel dieser dritten Studie ist es, mit wenig Aufwand eine Investitionsentscheidung zu ermöglichen: Soll *E-learning* in

den teilnehmenden Organisationen weiter ausgebaut werden? Schwerpunkt dieser Evaluationsstudie ist die Entwicklung von Methoden zur Ressourcenkontrolle als Evaluationsbaustein.

2.4 Kurzevaluation: *E-learning*

2.4.1 Hintergrund und Zielsetzung

E-learning wird in Organisationen zunehmend eingesetzt, um das durch Lehrer, *Coaches*, Instruktoren und andere Berufsgruppen vermittelte und fremdgesteuerte Lernen zu ergänzen. Die Lernenden haben dabei die Möglichkeit, durch den selbstverantwortlichen Umgang mit Computerprogrammen ihren Lernprozess selbst zu steuern. Das Lernen erfolgt auf dem *personal computer (computer based training)* oder aber auch im Internet (*web based training*). Es ist damit zeit- und raumunabhängig. *Web based training* ermöglicht darüber hinaus das Lernen in Gruppen und in verschiedenen Rollen (z.B. Piloten und Fluglotsen, die Fliegen und Flugsicherung simulieren). Jeder Schüler kann seinen individuellen Lernvorlieben und -fortschritten entsprechend vorgehen. Andererseits birgt *E-learning* die große Gefahr, dass Schüler mit mangelnden Fähigkeiten zur Selbstorganisation, Selbstdisziplinierung und Selbstmotivierung keinen Lernfortschritt erzielen.

Auf der Kostenseite hat *E-learning* den Nachteil, dass zunächst viele finanzielle und *Know-how*-Ressourcen in die Entwicklung der *E-learning* Instrumente investiert werden muss. Nach ihrer Entwicklung hat ein *E-learning* Instrument den außergewöhnlichen Vorteil, dass es einer beliebigen Menge an Schülern ohne nennenswerte weitere Kosten zugänglich gemacht werden kann. Sie sind darüber hinaus mit anderen *E-learning* Instrumenten kombinierbar und werden daher auch Module genannt.

Die Herausforderungen des *E-learning*s liegen - neben der anspruchsvollen technischen, grafischen und interaktiven Umsetzung - in der effektiven und effizienten Kombination mit anderen Lernformen, um die spezifischen Vorteile der jeweiligen Lernform zu nutzen und die spezifischen Nachteile zu kompensieren.

Diese Evaluationsstudie deckt nach HF und HR den dritten Zielbereich, Training, ab. Da *E-learning* in der Erstausbildung von Fluglotsen, dem so genannten *ab initio training*, untersucht wird, besteht im Gegensatz zu den bereits berufstätigen Fluglotsen oder *Call Center* Agenten keine direkte Wertschöpfung, die ökonomisch beziffert werden könnte. Daher erfolgt eine Bewertung der Kosteneffizienz mit Bezug auf die durch *E-learning* ersetzten klassischen Trainingsformen Frontal- und Simulatorunterricht. Wie in den anderen beiden Studien wurden Fähigkeiten, Arbeitstätigkeiten und Arbeitsergebnisse und deren Veränderung durch die Maßnahme, hier *E-learning*, erfragt. Leider wurden die Veränderungssitems nur von wenigen Auszubildenden beantwortet, so dass die Funktionsanalyse und die psychologische Wirkungsanalyse des HPM (Abbildung 1) nicht vollzogen werden konnten. Eine Organisations- und Prozessanalyse erfolgte ebenfalls nicht; diese Schritte waren im vorliegenden Anwendungsfall nicht notwendig, um eine Investitionsentscheidung über den zukünftigen Einsatz von *E-learning* zu fällen. Diese erforderte lediglich eine Beurteilung der Kosteneffizienz von *E-learning* im Vergleich zu den anderen Trainingsmethoden und damit den HPM-Schritt „Ressourcenkontrolle“. Eine Investitionsentscheidung mit so wenig Aufwand wie möglich war das übergeordnete Ziel dieser Kurzevaluation.

2.4.2 Untersuchungspersonen und –methoden

Wie in den anderen beiden Evaluationen wurden Daten zum *E-learning* von der Kernzielgruppe (in diesem Falle Flugsicherungsschüler, *Trainees* genannt), Fach-

und Führungskräften erhoben. Zunächst wurde der verantwortliche Projektmanager im Trainingsinstitut von EUROCONTROL, dem *Institute of Air Navigation Services* IANS in Luxemburg, über den derzeitigen Stand und die zukünftigen Entwicklungslinien des *E-learning*s in den europäischen Flugsicherungsakademien befragt. IANS entwickelt *E-learning* Module für die zu EUROCONTROL gehörenden Mitgliedstaaten, deren Flugsicherungsakademien die Module kostenlos nutzen dürfen. Einige Länder, wie z.B. Schweden, entwickeln auch eigene Module.

Zwei Flugsicherungsakademien, die in Schweden und die in Belgien, konnten für die Teilnahme an der Untersuchung gewonnen werden. Deren Manager und Lehrer wurden in einem ersten Schritt befragt (sieben Interviewpartner). Inhalte der Gespräche waren die aktuellen und zukünftigen Herausforderungen der Organisationen im Allgemeinen und Ziele, Inhalte, Umfang, Kosten und Nutzen der *E-learning* Module im Besonderen. Der Interviewleitfaden ist im Anhang wiedergegeben (Anh. 6: *E-learning* Interviewleitfaden Manager).

In einem zweiten Schritt nahmen 32 *Trainees* (13 in Schweden, 19 in Belgien) an einer Fragebogenstudie teil. Die Struktur des Fragebogens war wie folgt (s. auch Anh. 7: *E-learning* Traineefragebogen):

1. Selbsteinschätzung von 12 für die (zukünftige) Lotsentätigkeit wichtigen Fähigkeiten, die trainierbar sind, z.B. Wahrnehmungsgeschwindigkeit.
2. Selbsteinschätzung von 14 für die (zukünftige) Lotsentätigkeit wichtigen Tätigkeiten, z.B. Verkehrsbild beobachten und bewerten.
3. Selbsteinschätzung von 7 Ergebnisgrößen der Lotsentätigkeit z.B. Beitrag zur Luftverkehrssicherheit.

Diese ersten drei Variablengruppen entsprechen den nach HPM (Kapitel 1.4) zu berücksichtigenden Bereichen, wie sie zuerst im psycho-ökonomischen Rubikon-Modell der CISM-Studie dargestellt wurden (Abbildung 5). Positive Effekte von *E-learning* auf diese Fähigkeiten, Tätigkeiten und Ergebnisgrößen sollten im Fragebogen mit (++) oder (+) und keine oder negative Effekte mit (-) angezeigt werden. Leider machten nur sehr wenige *Trainees* von dieser Möglichkeit Gebrauch, so dass diese Daten nicht sinnvoll ausgewertet werden konnten.

Der nächste Abschnitt des Fragebogens enthielt eine

4. Liste mit *E-learning* Modulen, z.B. *Handover takeover*. Die *Trainees* waren hier gebeten anzukreuzen, mit welchen Modulen sie gearbeitet haben, wie viel Zeit sie damit zugebracht haben und welchen prozentualen Nutzen sie dem Modul zuweisen.
5. Aus dieser Liste sollte dann das Modul mit dem größten Nutzen näher bewertet werden. Ursprünglich war auch beabsichtigt, das schlechteste Modul in gleicher Weise bewerten zu lassen, um dann einen Durchschnittswert ermitteln zu können; diese Verdopplung des Befragungsaufwandes wurde von den Akademien nicht akzeptiert und daher fallen gelassen. In Belgien wurde jedoch zusätzlich ein neu eingeführtes Modul (*Handover takeover*) in gleicher Weise wie das Modul mit dem größten Nutzen bewertet:
 - a) Effekte auf Fähigkeiten, Tätigkeiten und Ergebnisgrößen wurden wie oben

- beschrieben leider kaum angegeben.
- b) Offene Antworten zu spezifischen Stärken und Schwächen des Moduls.
 - c) Anzahl der Stunden im Klassenzimmer und im Flugsicherungssimulator, die das Modul ersetzen konnte.
6. Es folgte eine generelle Bewertung des *E-learning* Programms in der jeweiligen Akademie:
- a) Wie viel Zeit verbrachte der *Trainee* mit Klassenunterricht, Simulator und *E-learning*.
 - b) Wie viel prozentualen Lerneffekt hatten diese verschiedenen Lernformen.
 - c) Wie viel der gesamten Lernzeit sollte in Zukunft auf *E-learning* verwendet werden.
 - d) Offene Antworten zu allgemeinen Stärken und Schwächen des *E-learning*s.
 - e) Prozentuale Angaben zur Erfüllung guter Gestaltungskriterien der *E-learning* Module, z.B. Navigierbarkeit, Handhabung.

Wie in den anderen beiden Evaluationen wurde zur Validierung der Ergebnisse außer der Sichtweise der Kernzielgruppe auch eine Fremdsicht erhoben. In diesem Falle waren die Auskunftgeber vier Lehrer an den Akademien, im Folgenden *Trainer* genannt. Sie beantworteten dieselben Fragen in Bezug auf ihre *Trainees*, z.B. schätzten sie die Fähigkeiten, Tätigkeiten und Ergebnisgrößen für einen durchschnittlichen *Trainee* ein und wie diese durch *E-learning* verändert werden (Anh. 8: *E-learning Trainerfragebogen*).

2.4.3 Ergebnisse

Die im vorangegangenen Kapitel beschriebenen Fragebogen wurden an die Lotsenschülerinnen und -schüler (im Folgenden *Trainees*) der Akademie von Belgocontrol (am Flughafen Brüssel) und Luftfartsverket (am Flughafen Sturup in Schweden) verteilt. In Schweden werden derzeit 29 schwedische und norwegische *Trainees* ausgebildet, da die skandinavischen Länder vor kurzem ihre Akademien in Malmö zusammengelegt haben. Dänemark bildet derzeit nicht aus. Um die Lesbarkeit dieses Berichts zu erhöhen, wird im Folgenden vereinfachend von der „schwedischen Stichprobe“ oder den „schwedischen *Trainees*“ geschrieben. Von den 29 *Trainees* füllten 13 den Fragebogen aus, die Rücklaufquote beträgt also knapp 50% und ist zufriedenstellend. In Belgien werden derzeit 24 *Trainees* ausgebildet, von denen 19 ihren Fragebogen zurücksandten, die Rücklaufquote ist mit knapp 60% sehr gut. Um die Antworten der *Trainees* mit denen ausgebildeter Fluglotsen zu vergleichen, wurden die im Rahmen der CISM-Evaluation erhobenen Fähigkeiten, Arbeitstätigkeiten und Leistungen der 20 jüngsten und hinsichtlich Geschlecht parallelierten Fluglotsen herangezogen (Kapitel 2.2). Es nahmen auch vier Flugsicherungslehrer (im Folgenden *Trainer*) an der Untersuchung teil, indem sie eine Fremdeinschätzung von Fähigkeiten, Arbeitstätigkeiten und Leistungen ihrer *Trainees* abgaben und wie diese ggf. durch *E-learning* beeinflusst werden. Aufgrund ihrer geringen Zahl wird auf die *Trainer*-bogen nur am Rande eingegangen. Das Hauptziel dieser Studie war es, die Effektivität sowie Effizienz der *E-learning* Module im Paretoprinzip nur mit so viel Aufwand zu überprüfen, dass eine Managemententscheidung zum *E-learning* getroffen werden kann.

2.4.3.1 Fähigkeiten, Arbeitstätigkeiten und Leistung

Wie bei den beiden anderen Evaluationen werden zunächst Fähigkeiten, Arbeitstätigkeiten und Leistungen der *Trainees* dargestellt.

Die Daten wurden durch eine generelle Selbsteinschätzung der beobachteten Personen mit einer prozentualen Skala von 0-100% erhoben. Um die genauen Items zu generieren, wurden die Evaluationsindikatoren herangezogen (Kapitel 2.4.1).

Das daraus entstehende Stärken-Schwächen-Profil der Befragten ist in Abbildung 51 dargestellt. Abbildung 51 zeigt, dass Fluglotsen und *Trainees* die verschiedenen Fähigkeiten bei sich zu etwa 70% ausgeprägt berichteten. Die Fluglotsen bewerteten einige Fähigkeiten wie etwa selektive Aufmerksamkeit, *Multitasking* und Teamarbeit höher als die *Trainees*. Dies war zu erwarten, denn die *Trainees* hatten ihre Ausbildung noch nicht abgeschlossen und noch keine Gelegenheit, ihre ATC-Fähigkeiten durch Arbeitsroutine zu vertiefen. Die schwedischen *Trainees* schätzten ihre Fähigkeiten geringer ein als die belgischen, z.B. hinsichtlich ihrer „*self-awareness*“. Die Länderunterschiede können auf zwei verschiedene Weisen erklärt werden:

1. Kann ein realer Unterschied zwischen den Fähigkeiten der *Trainees* beider Länder vorliegen. Diese Erklärung ist unwahrscheinlich, da die Selektionsverfahren in beiden Ländern gleich sind.
2. Ist eine gewisse skandinavische Bescheidenheit in vielen kulturellen Kontexten bekannt (z. B. Vogt et al., im Druck). Es ist daher wahrscheinlich, dass sich auch die in dieser Kultur sozialisierten Flugsicherungs*trainees* entsprechend verhalten.

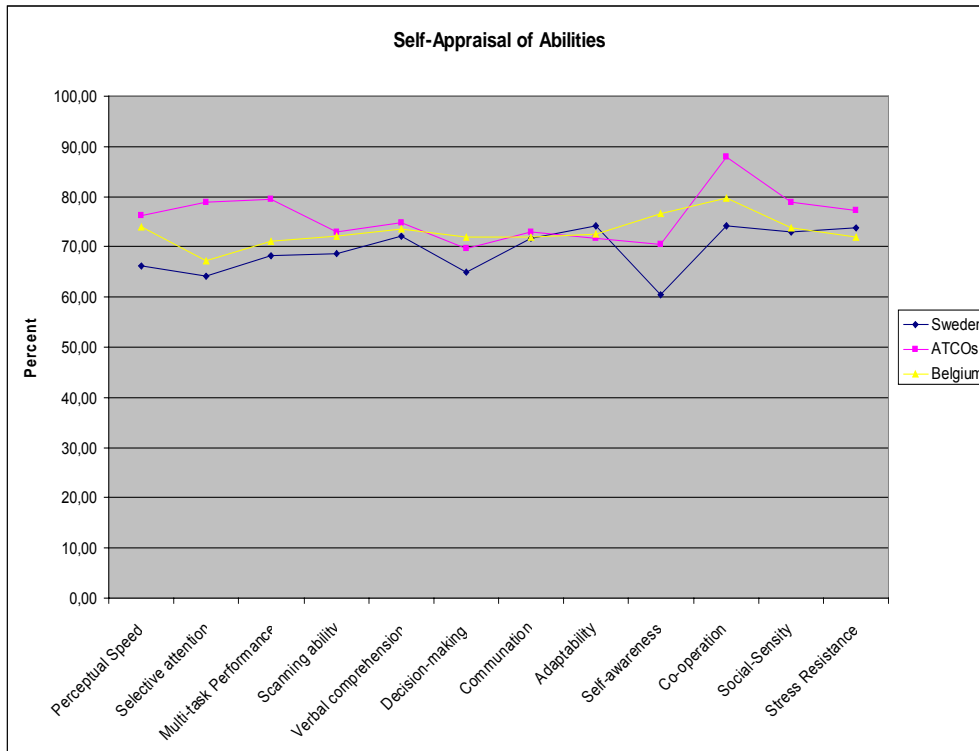


Abbildung 51: Mittelwerte selbst berichteter Fähigkeiten in Prozent für 13 schwedische und 19 belgische *Trainees* sowie 20 ausgebildeter Fluglotsen (ATCO)

Abbildung 52 zeigt die berichtete Qualität der Arbeitstätigkeiten. Die Bewertungen der *Trainees* beziehen sich dabei überwiegend auf ihre Arbeit im Flugsicherungssimulator. Die Abbildung zeigt, dass die Qualität der Arbeitstätigkeiten wiederum von allen Gruppen bei 70-80% gesehen wurde. Im Gegensatz zu den Fähigkeiten schätzten *Trainees* und zertifizierte Fluglotsen die meisten Arbeitstätigkeiten jedoch gleich ein. Die Überlegenheit von 5 bis 10% zu Gunsten der Fluglotsen taucht nur noch bei Tätigkeiten auf, die im Arbeitsalltag der Fluglotsen routinisiert werden, wie etwa Piloten instruieren, Flugstreifen bearbeiten und Kollegen unterstützen.

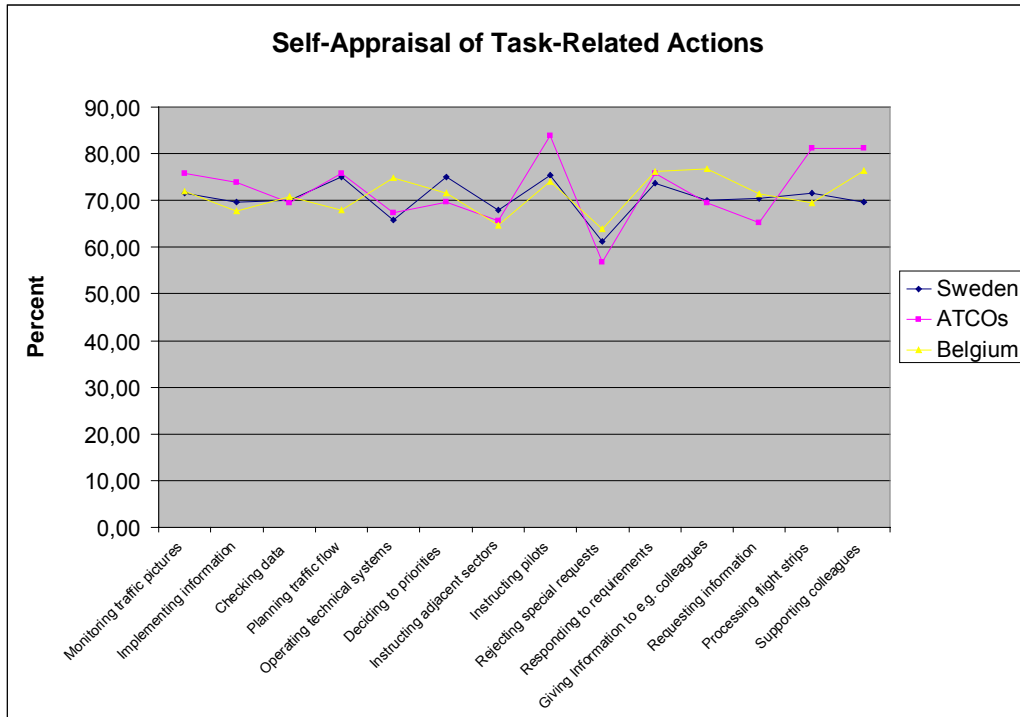


Abbildung 52: Mittlere selbst berichtete Qualität der Arbeitstätigkeiten in Prozent für 13 schwedische und 19 belgische *Trainees* sowie 20 ausgebildeter Fluglotsen (ATCO)

Abbildung 53 zeigt die Bewertung von Arbeitsleistungen. Die Fluglotsen gaben ihre Leistungen im Durchschnitt mit 80% an, die *Trainees* mit 70%. Der Erfahrungsvorteil der Fluglotsen kommt also ebenso zum Tragen wie bei den Fähigkeiten. Der konsistente Unterschied zu Gunsten der belgischen *Trainees* blieb bei den Arbeitsleistungen jedoch aus.

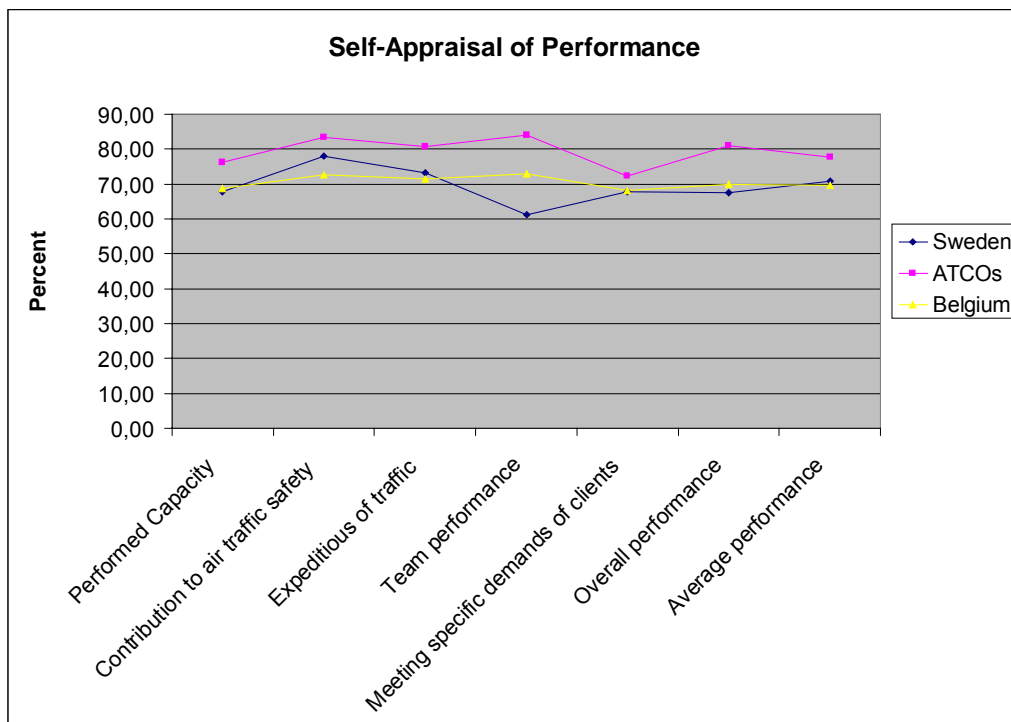


Abbildung 53: Mittlere selbst berichtete Leistung in Prozent für 13 schwedische und 19 belgische *Trainees* sowie 20 ausgebildeter Fluglotsen (ATCO)

2.4.3.2 *E-learning* im Vergleich zu Simulator und Klassenunterricht

E-learning wird mit dem Ziel eingesetzt, die bisherigen klassischen Unterrichtsformen im Klassenraum und im Simulator zu ersetzen bzw. zu ergänzen. Ein Effektivitäts- und Effizienzvergleich der drei Lernformen liegt daher nahe. Im Folgenden wird dargestellt, wie viel Zeit die *Trainees* mit der jeweiligen Lernmethode verbringen und wie effektiv bzw. effizient sie bewertet wird.

Abbildung 54 zeigt die investierte Zeit und den empfundenen Lerneffekt für die schwedische Datenbasis. Deutlich wird auf den ersten Blick, dass die *Trainees* mit 55% die meiste Zeit im Simulator verbrachten und dass hier auch der größte Lerneffekt gesehen wurde (67%). Auf das Klassenzimmer entfielen 36% der Lernzeit und 29% des Lerneffekts, d.h. die *Input-Output*-Relation ist negativ. Dies ist auch der Fall für *E-learning*, für das die *Trainees* 9% ihrer Zeit aufwendeten und einen Lerneffekt von nur 4% berichteten. Dies deutet zum einen auf mögliche qualitative Verbesserungspotentiale der *E-learning* Module hin, auf die an späterer Stelle noch detaillierter eingegangen wird (Kapitel 0). Zum anderen kann vermutet werden, dass die schwedischen *Trainees* aufgrund der langen Tradition dieser Lernform in ihrer Organisation, die auch eigene Module passgenau entwickelt, höhere Ansprüche an das *E-learning* stellen und daher in ihrem Urteil kritischer sind.

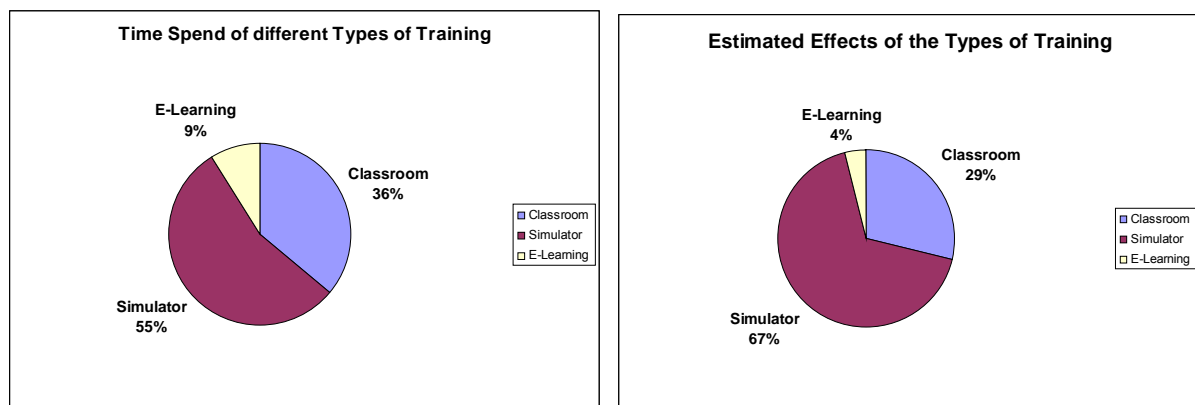


Abbildung 54: Links: Mittlere Zeit in Prozent, die schwedische *Trainees* auf *E-learning*, Klassenunterricht und Simulatortraining verwendeten

Rechts: Mittlerer Lerneffekt (Nutzen) in Prozent, den die schwedischen *Trainees* *E-learning*, Klassenunterricht und Simulatortraining zuwiesen

Um die *Input-Output*-Relationen in einer einzelnen Grafik darzustellen, wurde der Quotient (Lerneffekt pro Zeit) gebildet (Abbildung 55). Ein Quotient von über 1 für das Simulatortraining drückt wiederum aus, dass diese Lernmethode mehr Lerneffekt bringt als sie Zeit kostet. Klassenunterricht und *E-learning* erreichen dagegen nur *Input-Output*-Relationen von unter 1.

Die Standardfehler (SE für *Standard Error*) sind bei den Quotienten relativ klein, d. h. die *Trainees* waren in ihren Urteilen einig.

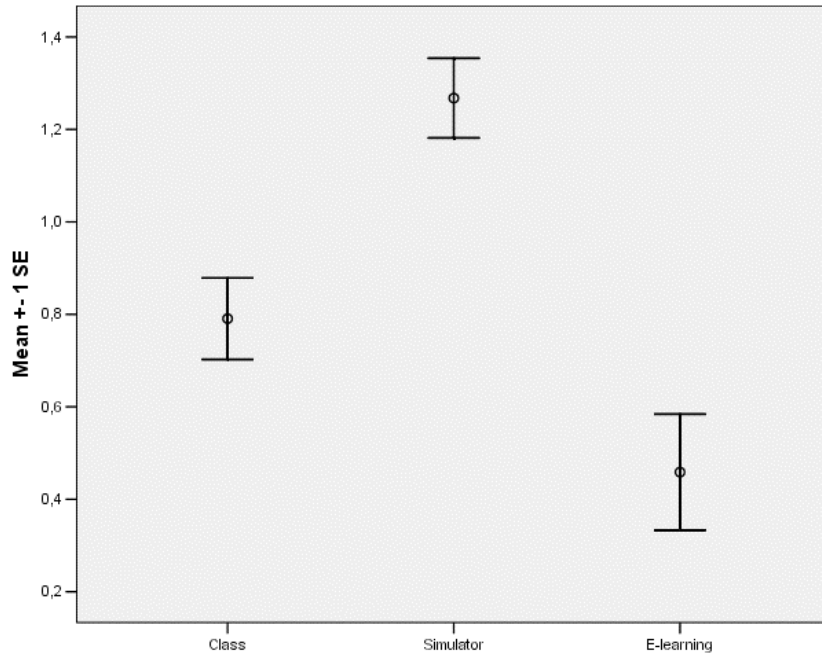


Abbildung 55: Mittlerer Quotient und Standardfehler von wahrgenommenem Lerneffekt pro eingesetzte Zeit in der schwedischen Stichprobe

Die gleichen Auswertungen wurden für die belgische Stichprobe vorgenommen. Hier wurden 48% der Lernzeit für den Simulator berichtet. Mit 31% berichteten die belgischen *Trainees* 5% weniger Klassenunterricht als die Schweden. Diese 5% verbrachten die Belgier nach eigenen Angaben zusätzlich mit *E-learning*. Dieses Ergebnis ist insofern überraschend, als die Skandinavier auf eine längere *E-learning* Tradition zurückschauen können.

Bei der berichteten Effektivität von *E-learning* waren die Urteile der belgischen *Trainees* etwas positiver: 10% Lerneffekt standen 21% investierter Lernzeit gegenüber und resultierten in einer *Input-Output*-Relation von etwa 0,5. Auch die Effizienz des Klassenunterrichts sahen die Belgier tendenziell positiver, sie gaben eine *Input-Output*-Relation von 0,8 an. Dem Simulator schrieben auch die Belgier mit 48% investierter Zeit und 64% erlebten Nutzen die beste *Input-Output*-Relation zu, sie beurteilten die Effizienz des Simulators wiederum etwas besser als die Schweden.

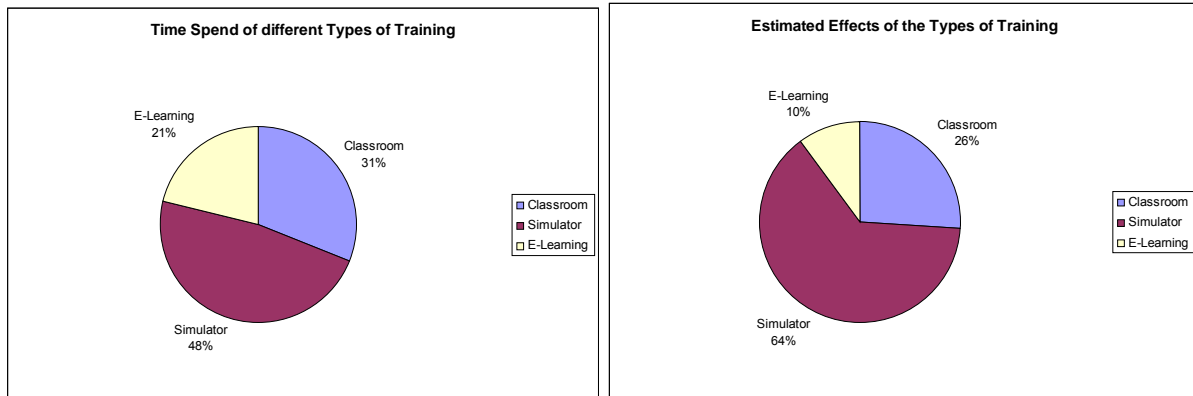


Abbildung 56: Links: Mittlere Zeit in Prozent, die belgische *Trainees* auf *E-learning*, Klassenunterricht und Simulatortraining verwendeten
Rechts: Mittlerer Lerneffekt (Nutzen) in Prozent, den die belgischen *Trainees* *E-learning*, Klassenunterricht und Simulatortraining zuwiesen

Kurz zusammengefasst konnte in diesem Abschnitt gezeigt werden, dass beide Stichproben *E-learning* einen – wenn auch geringen – Nutzen zuschreiben. In diesem Sinne ist *E-learning* zwar effektiv jedoch nicht effizient in Bezug auf die investierte Zeit. Im nächsten Schritt wird die ökonomische Effizienz betrachtet. Obwohl *E-learning* weniger effizient bezüglich Lerneffekt pro Zeit war, wird sich zeigen, dass es ökonomisch sinnvoll ist, da es gegenüber Klassenunterricht und Simulator Kostenvorteile hat, die deren Effizienzvorteil überkompensieren.

2.4.3.3 Monetarisierung von E-learning als Ersatz für Simulator/Klasse

Die schwedischen *Trainees* wurden gebeten, für das ihrer Meinung nach beste *E-learning* Modul anzugeben, wie viele Stunden Simulatortraining oder Klassenunterricht es bei ihnen ersetzt hat. Die Darstellung erfolgt getrennt für unterschiedliche *E-learning* Module, weil die *Trainees* unterschiedliche Module an die erste Stelle setzten. Dies waren *Arriving IFR*⁹ (*Instrumental Flight Routines*, 1 *Trainee*), *Ground lights*¹⁰ (4 *Trainees*), *Holding pattern*¹¹ (1 *Trainee*) und *Radar theory*¹² (3 *Trainees*). Bei 4 Personen fehlten leider Angaben zum besten Modul.

Abbildung 57 zeigt, dass die verschiedenen *E-learning* Module zwischen knapp einer und vier Stunden Klassenunterricht ersetzen konnten. Das Modul *Arriving IFR* ersetzte eine Stunde Klassenunterricht bei einer Person, *Ground lights* ersetzte im Mittel eine Dreiviertelstunde bei vier *Trainees* und *Holding pattern* sowie *Radar theory* vier Stunden bei einer bzw. bei drei Personen. Letztere waren sich bei dieser Einschätzung sehr uneinig, wie der große Standardfehler des Mittelwerts zeigt.

⁹ Dieses *E-learning* Modul trainiert die Koordination ankommender Flugzeuge, i.d.R. alle Passagiermaschinen, die einen Flughafen mit Hilfe technischer Systeme wie Leitstrahlen (im Gegensatz zum Sichtflug) anfliegen.

¹⁰ Dieses *E-learning* Modul trainiert die Berücksichtigung der verschiedenen Leuchtsignale, die an Flugzeugen und am Flughafen eingesetzt werden.

¹¹ Dieses *E-learning* Modul trainiert die Koordination von Flugzeugen in Warteschleifen (*holdings*)

¹² Dieses *E-learning* Modul trainiert die Grundlagen der Flugsicherung: Wie funktioniert die Radarerkennung und –darstellung von Flugzeugen, wie werden Flugdaten verarbeitet etc.

Simulatorstunden konnten durch *E-learning* nicht in diesem Umfang ersetzt werden. Nur 1,5 Simulatorstunden wurden im Mittel als durch das Modul *Radar theory* eingespart berichtet. Die Streuung der Urteile war noch größer als bei der Frage nach ersetzttem Klassenunterricht.

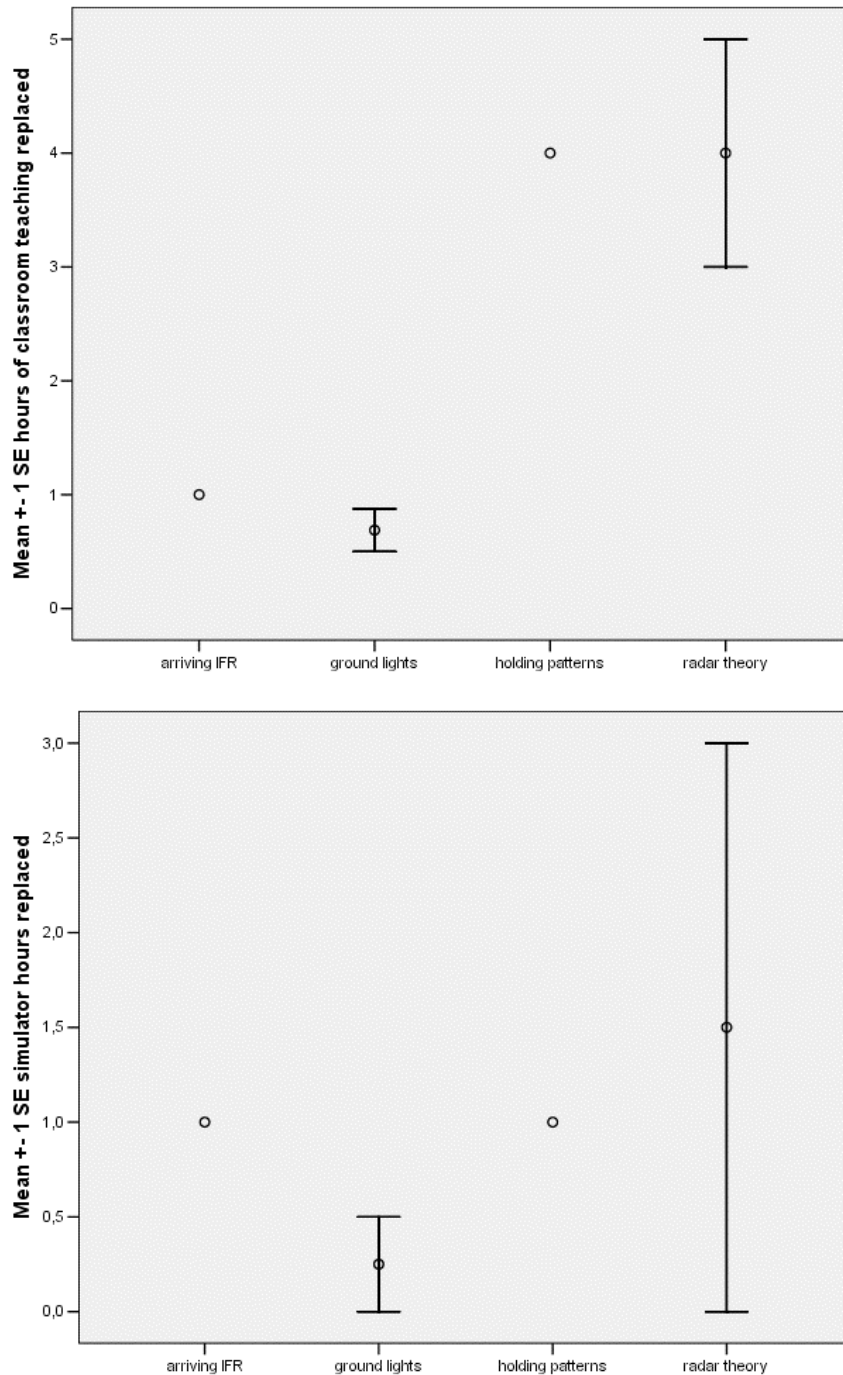


Abbildung 57: Mittlere Anzahl Klassen- (oben) und Simulatorstunden (unten), die *E-learning* in der schwedischen Stichprobe ersetzt hat

Als Fazit dieser Auswertung kann festgehalten werden, dass *E-learning* andere Unterrichtstypen ersetzen kann. Allerdings wurde in den offenen Antworten ein Verbesserungsbedarf bei der Anpassung der Module an die lokalen Gegebenheiten wie z. B. belgische oder schwedische Flughäfen im Modul *Radar theory* gesehen, wodurch die Effektivität der Kurse erhöht werden könnte.

2.4.3.4 Return on Investment und Break Even der E-learning Nutzung

Die dritte Ebene der Evaluation beschäftigt sich mit der ökonomischen Effizienz der Implementierung von *E-learning*. Tabelle 2 zeigt den ROI der Investitionen der unterschiedlichen *E-learning* Module. Die Kalkulation beruht auf dem Produkt der Anzahl der schwedischen *Trainees* mit der durchschnittlichen Zahl der Simulator- und Klassenstunden, die in der schwedischen Stichprobe ersetzt wurde. Das Rationale für dieses Vorgehen ist, dass grundsätzlich alle Module allen *Trainees* jederzeit zur Verfügung stehen. Es nahmen jedoch nicht alle *Trainees* an der Untersuchung teil und nicht alle Teilnehmer bewerteten ein- und dasselbe Modul als das beste. Die ersetzte Zeit im Klassenzimmer bzw. im Simulator wurde daher auf der Grundlage der vorhandenen Daten für alle *Trainees* hochgerechnet. Dieses Vorgehen ist nicht wie bisher konservativ und führt daher eher zu **potenziellen** als realen ROI.

Da *E-learning* Module relativ teuer in der Entwicklung sind, jedoch nur wenige Applikationskosten verursachen, ist es wahrscheinlich, dass die ermittelten potenziellen ROI auch erreicht werden. Dies ist nur eine Frage einer bestimmten Anzahl von *Trainees*. Die untersuchten Akademien sind relativ klein, so dass die ermittelten ROI der *E-learning* Programme wiederum Unterschätzungen des tatsächlichen Potenzials darstellen. Größere Flugsicherungsakademien oder gar IANS, das für alle europäischen Flugsicherungen *E-learning* entwickelt, erreichen die ökonomische Effizienz in wesentlich kürzeren Zeiträumen.

Zunächst wurde der *Return on Investment* (ROI) berechnet nach folgender Formel:

$$100 * \left(\frac{1 + \text{Ersparnis ersetzen Unterrichts Klasse und Simulator} - \text{Kosten des Moduls}}{\text{Kosten des Moduls}} \right)$$

Außer dem ROI für jedes einzelne *E-learning* Modul wurde die Anzahl von *Trainees* in den Akademien ermittelt, ab der sich die Investition in dieses Modul rentiert hat. Dieser Amortisationszeitpunkt (*Break Even Point*, BEP) ermöglicht es den Managern der Akademien, den Zeitraum zu überschauen, in dem die Rückflüsse durch das Modul die Investition unterschreiten. Sie können so absehen, ob sich die Investition in ein oder zwei Jahren rentiert haben wird, wenn weitere *Traineejahrgänge* das Modul nutzen konnten. Der BEP wurde nach folgender Formel berechnet:

$$\frac{\text{Kosten des Moduls}}{\text{Kostensparnis durch ersetzen Unterricht / Trainee}}$$

Da die schwedische Akademie mehrere der betrachteten *E-learning* Module selbst entwickelt hat, konnten die Vollkosten genau ermittelt werden. Sie sind mit 4.000 bis 5.000€ vergleichsweise niedrig, vermutlich weil eine jahrzehntelange Erfahrung in der Organisation vorliegt. Der ökonomische Nutzen der schwedischen Module wurde über die Kosten der durchschnittlich ersetzten Simulator- und Klassenstunden bestimmt. Eine Simulatorstunde kostet nach Auskunft der schwedischen Akademie 120€. Für eine Stunde Klassenunterricht fällt lediglich eine *Trainerstunde* in Höhe von 60€ an.

Die Nutzung von EUROCONTROL Modulen ist für die Mitgliedsstaaten in den Gebühren, die sie an den Brüsseler Dachverband zahlen, enthalten. Entsprechend schwierig gestaltet sich die Ermittlung der Kosten, die der einzelnen Akademie durch *E-learning* entstehen. Es wurden daher die Kosten zugrunde gelegt, die IANS für die

Entwicklung der Module ansetzt. Sie liegen zwischen 6.000 und 15.000€. Für das in Schweden genutzte Modul *Radar theory* veranschlagte IANS Vollkosten i.H.v. 15.000€, die für die ROI-Berechnung herangezogen wurden. Dies entspricht wiederum einer konservativen Annahme, nämlich dass die Akademie in Schweden alleine die gesamten Entwicklungskosten für *Radar theory* zu tragen hätte.

Tabelle 2 zeigt für die vier betrachteten *E-learning* Module, wie viele Stunden herkömmlichen Unterrichts im Mittel ersetzt wurden. Die Kosten, die durch den herkömmlichen Unterricht entstanden wären, wurden durch die o.g. Stundensätze ermittelt und als durchschnittlicher monetärer ROI pro *Trainee* interpretiert. Die Akademie in Malmö bildet derzeit 29 *Trainees* aus, die alle Zugang zu den *E-learning* Modulen haben. Der ROI pro *Trainee* wurde daher mit 29 multipliziert. Das Produkt wurde als Gesamt-ROI des *E-learning* Programms der schwedischen Akademie gewertet, in der Annahme, dass die Angaben in unserer schwedischen Stichprobe repräsentativ sind und auch für die *Trainees* gelten, die nicht an der Studie teilgenommen haben.

Das Modul *Radar theory* ersetzt die meisten Simulator- und Unterrichtsstunden. Diese hohe Kostenersparnis muss jedoch relativiert werden an den hohen Kosten des Moduls, der ROI liegt daher noch knapp unter den Entwicklungskosten. Der BEP beträgt 36 *Trainees*, d.h. mit der Aufnahme von nur sieben weiteren *Trainees* wird der ROI 100% überschreiten und das Modul hat sich amortisiert.

Der ROI ist mit 261% für das Modul *Holding patterns* am höchsten, weil dieses Modul eine kostengünstige Eigenentwicklung der schwedischen Akademie war und gleichzeitig fünf Stunden der anderen Lernformen ersetzen kann. Dieses Modul hatte sich bereits nach fünf Nutzern vollständig bezahlt gemacht.

Für das Modul *Arriving IFR* beträgt der ROI 130%. Die Investition in die Entwicklung dieses Moduls hatte sich nach 22 Nutzern rentiert.

Die Amortisierung des Moduls *Ground lights* lässt am längsten auf sich warten. Dies ist sowohl dem geringen Nutzen (weniger als eine Stunde herkömmlichen Unterrichts wurde ersetzt) als auch den relativ hohen Kosten zuzuschreiben. Der ROI liegt bei 31%, der BEP bei 93 *Trainees*. Es werden also mindestens weitere zwei Jahrgänge gleicher Stärke (29 *Trainees*) benötigt, bis sich die Investition in dieses Modul beginnt auszuzahlen.

Tabelle 2: Mittlere Anzahl M von Simulator- und Klassenstunden, die durch die besten *E-learning* Module ersetzt wurden (Angaben von 8 schwedischen *Trainees* bei 29 *Trainees* insgesamt).

| | <i>Arriving IFR</i> | <i>Ground lights</i> | <i>Holding-patterns</i> | <i>Radar theory</i> |
|------------------------------|---------------------|----------------------|-------------------------|---------------------|
| <i>M class replaced</i> | 1 | 0,7 | 4 | 4 |
| <i>M sim. replaced</i> | 1 | 0,1 | 1 | 1,5 |
| <i>Revenue / Trainee</i> | 180 | 54 | 360 | 420 |
| <i>Revenue / 29 Trainees</i> | 5.220 | 1.566 | 10.440 | 12.180 |
| <i>Costs for LV</i> | 4.000 | 5.000 | 4.000 | 15.000 |
| <i>ROI / 29 Trainees</i> | 131% | 31% | 261% | 81% |
| <i>BEP in Trainees</i> | 22 | 93 | 11 | 36 |

Bei der Kalkulation für den *Return on Investment* der belgischen Akademie sind wir in gleicher Weise vorgegangen. Die Anzahl der *Trainees* hier beträgt 24. Kosten für Klassenunterricht und Simulatorstunden fallen in Höhe von 24,5 Euro bzw. 332 Euro (Simulator) an. Die Unterschiede in den schwedischen und belgischen Kostensätzen sind bedeutsam, konnten bisher jedoch nicht aufgeklärt werden.

Belgocontrol verwendet ausschließlich von EUROCONTROL entwickelte *E-learning* Module, so dass wie im Falle des in Schweden genutzten EUROCONTROL Moduls *Radar theory* vorgegangen wurde. Die von IANS angegebenen Entwicklungskosten sind in Tabelle 3 angegeben.

Das Modul *Handover takeover*¹³ war zum Zeitpunkt der Datenerhebung (März, April und Juli 2006) relativ neu (Einführung Februar 2006). Da hier die Einführung begleitet werden konnte, wurde dieses Modul in jedem Fragebogen behandelt. Allerdings hatte die Mehrzahl der *Trainees*, die im März und April befragt wurde, nicht genügend Erfahrung, um die Nutzenangaben zum *Handover takeover* Modul auszufüllen. Acht *Trainees*, insbesondere die im Juli befragten, machten Nutzenangaben, d.h. sie berichteten, wie viele Stunden herkömmlichen Unterrichts ersetzt wurden. Tabelle 3 zeigt, dass sie von durchschnittlich 2,75 bzw. 1,25 gesparten Klassen- und Simulatorstunden berichteten, die einen ROI von 23% ergaben. Dieses Modul wird sich demnach erst in einigen Jahren rentiert haben.

Außer einer Bewertung des *Handover takeover* Moduls verlangte der Fragebogen die Bewertung eines weiteren Moduls. Hier sollten die *Trainees* das aus ihrer Sicht beste

¹³ Dieses *E-learning* Modul trainiert die Übergabe des Fluglotsenarbeitsplatzes bei Schichtwechsel oder Pausenablösung. Der Positionsinhaber macht dabei seinen Nachfolger z.B. auf kritische Wetter- und Verkehrssituationen aufmerksam.

Modul auswählen. Insgesamt machten weitere acht *Trainees* von dieser Möglichkeit Gebrauch.

Die Bilanz des Moduls *Runway incursion*¹⁴ war positiver als die des *Handover takeover* Moduls, insbesondere da ersteres deutlich weniger Kosten verursacht hat. Im Durchschnitt gaben die *Trainees* an, dass dieses Modul ihnen 12 Minuten Simulator- und zweieinhalb Stunden Klassenunterricht erspart hat. Der ROI von 51% zeigt an, dass die Hälfte der Entwicklungskosten für dieses Modul bereits in diesem Jahr zurückgeflossen ist. Das Modul wird sich im dritten Jahr rentiert haben, nachdem über 47 *Trainees* es genutzt haben (Tabelle 3).

Das Modul *Emergencies*¹⁵ konnte fast fünf Stunden Klassenunterricht und zwei Stunden Simulatorzeit ersetzen. Der ROI ist mit 120% sehr positiv. Das Modul hatte den BEP bereits nach 21 Nutzern erreicht.

Das Modul *Navigation*¹⁶ ersetzte dagegen Klassen- und Simulatorunterricht im gleichen Umfang. ROI und BEP zeigen, dass sich dieses Modul schon nahezu rentiert hat und trotz der hohen Kosten im nächsten Jahr schwarze Zahlen schreiben wird.

Tabelle 3: Mittlere Anzahl M von Simulator- und Klassenstunden, die durch die besten *E-learning* Module ersetzt wurden (Angaben von 16 belgischen *Trainees* bei angenommenen 24 *Trainees* insgesamt).

| | <i>Handover takeover</i> | <i>Emergencies</i> | <i>RWY incursion</i> | <i>Navigation</i> |
|------------------------------|--------------------------|--------------------|----------------------|-------------------|
| <i>M class replaced</i> | 2,75 | 4,8 | 2,5 | 1,5 |
| <i>M sim. replaced</i> | 1,25 | 1,9 | 0,2 | 1,5 |
| <i>Revenue / Trainee</i> | 481 | 748,4 | 127,7 | 534,8 |
| <i>Revenue / 29 Trainees</i> | 11.544 | 17.962 | 3.063,6 | 12.834 |
| <i>Costs for IANS</i> | 51.000 | 15.000 | 6.000 | 15.000 |
| <i>ROI / 24 Trainees</i> | 23% | 120% | 51% | 86% |
| <i>BEP in Trainees</i> | 107 | 21 | 47 | 29 |

¹⁴ Dieses *E-learning* Modul trainiert den Umgang mit Verkehrskonflikten auf der Start-/Landebahn (*runway*) eines Flughafens; diese können z.B. auftreten zwischen einem Flugzeug, das die Bahn kreuzt, um zur Parkposition zu rollen, und einer startenden oder landenden Maschine; aber auch Bodenfahrzeuge können in Konflikt mit Flugzeugbewegungen geraten.

¹⁵ Dieses *E-learning* Modul trainiert den Umgang mit Notfällen, z.B. technischen Ausfällen an Bord einer Maschine.

¹⁶ Dieses *E-learning* Modul trainiert alle Methoden der Navigation, z.B. das Berechnen eines Flugkurses oder das Lesen von Luftkarten.

2.4.3.5 Perspektiven der Verbesserung von *E-learning*

Die Rentabilität von *E-learning*, zumindest wenn das beste Modul bewertet wird, wurde in den vorangegangenen Abschnitten dargestellt. Die letzte Ebene der Analyse beschäftigt sich nun mit Fragen guter Gestaltung und eventuellem Handlungsbedarf. Die *Trainees* wurden gebeten, eine allgemeine Einschätzung des Nutzens der einzelnen Module und eine Bewertung hinsichtlich vorgegebener Gestaltungskriterien zu geben. Die schwedischen *Trainees* schätzten den prozentualen Nutzen der EUROCONTROL Module mit 40-50% ein, den der Module aus der eigenen Akademie etwas höher mit 50-60% (Abbildung 58). Dies entspricht dem Bild, dass sich die schwedischen *Trainees* durch ein höheres Anspruchsniveau und damit eine kritischere Haltung gegenüber *E-learning* auszeichnen.

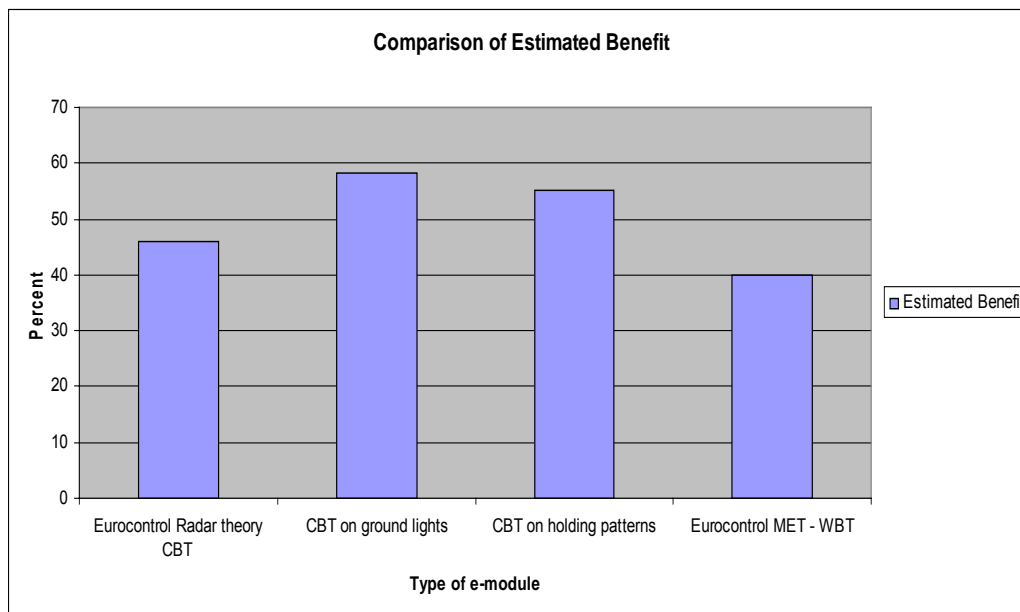


Abbildung 58: Wahrgenommener, mittlerer Nutzen der vier schwedischen *E-learning* Module

Abbildung 59 zeigt die Ergebnisse der Belgier. Der von den belgischen *Trainees* gesehene Nutzen lag um etwa 10% höher als bei den Schweden. Die etablierten Module *Emergency*, *Runway incursion* und *Navigation* erreichten Nutzenwerte von 60% oder mehr, das neue Modul *Handover takeover* ebenfalls.

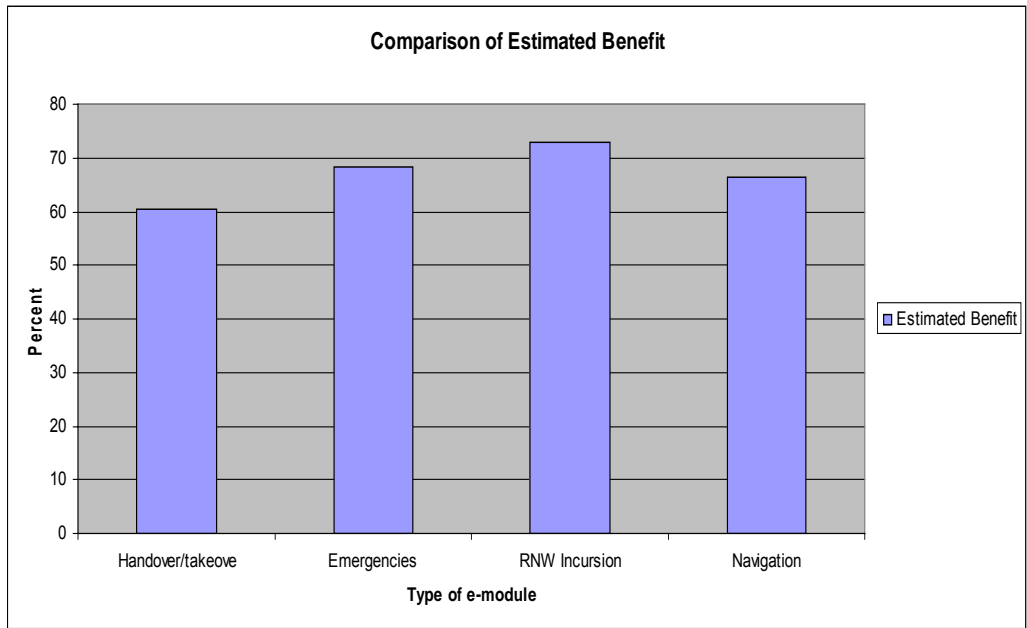


Abbildung 59 Wahrgenommener, mittlerer Nutzen der vier belgischen *E-learning* Module

Um festzustellen, wo die schwedischen und belgischen *Trainees* die nicht erfüllten Nutzenprozente sahen, wurden sie gebeten, das ihrer Meinung nach beste Modul hinsichtlich einer Liste von Gestaltungskriterien zu bewerten, die einer Literaturliteraturanalyse (Sars, 2006) entnommen wurden (Abbildung 60).

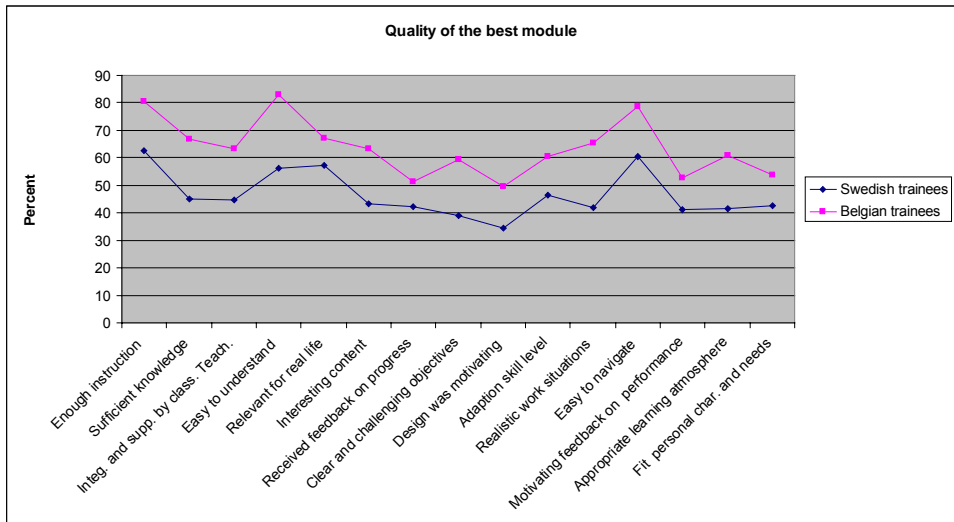


Abbildung 60: Prozent erfüllte Qualitätskriterien der schwedischen und belgischen *E-learning* Module

Abbildung 60 zeigt, dass die belgischen *Trainees* die guten Gestaltungskriterien zu 50-80% erfüllt sahen, während ihre schwedischen Kollegen dies nur zu 30-60% angaben. Die kritischere Haltung der schwedischen *Trainees* kam damit auch bei diesen Fragen zum Ausdruck. Beide *Trainee*gruppen kritisierten die mangelnde individuelle Anpassungsfähigkeit der Module. Auch das durch *E-learning* erworbene Wissen, die Integration in *Curriculum* und Kontext, interessante Gestaltung, mangelndes *Feedback*, Zielklarheit, Realitätsnähe und Lernatmosphäre wurden gering eingeschätzt. Die Motivation zum *E-learning* durch Zielklarheit, Einbettung in den Lehrplan, *Feedback* und eine gute Lernatmosphäre scheint ein zentraler Ansatzpunkt der *Trainer* zur Optimierung zu sein. Diesen Optimierungsbedarf hat IANS bereits vor einem Jahr aufgegriffen und eine *Workshop*reihe für *Trainer* begonnen, in der auf die Im-

plementierung und die Integration von *E-learning* in den Lehrplan fokussiert wird. Die technische Gestaltung dagegen wurde relativ gut beurteilt, d.h. die Module wurden als gut instruiert, gut verständlich und leicht zu navigieren beschrieben.

Abschließend sollten die *Trainees* einschätzen, wie sie persönlich von *E-learning* profitiert haben und wie viel Prozent der Ausbildung ihrer Meinung nach im *E-learning* vollzogen werden sollte (Abbildung 61). Wiederum lag der berichtete persönliche Nutzen der Schweden um 15% niedriger als der der Belgier. Die Schweden gaben im Durchschnitt an, dass ein Fünftel der Ausbildung mit *E-learning* bestritten werden sollte, die Belgier empfahlen sogar ein Viertel. Damit waren sich beide Gruppen darin einig, dass – bei aller Kritik – der bisherige (Kapitel 2.4.3.2) Zeitaufwand für *E-learning* erhöht werden sollte.

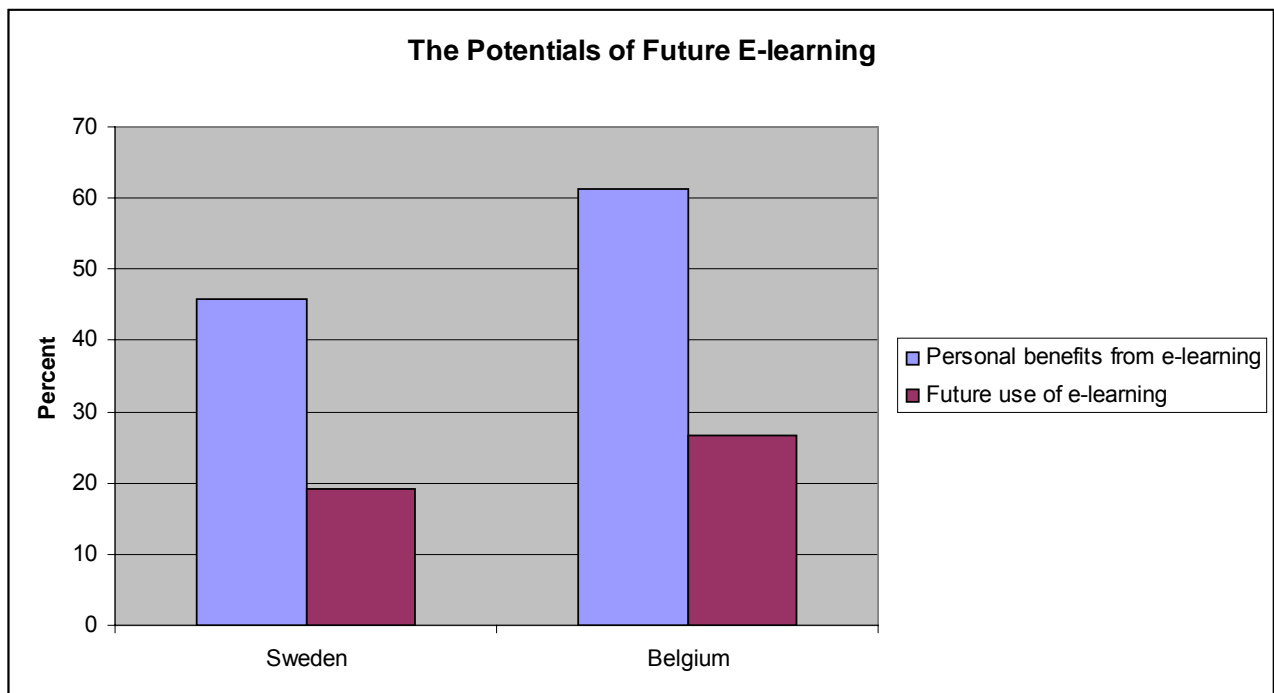


Abbildung 61: Persönlicher Nutzen und empfohlener Ausbildungsanteil von *E-learning* in Schweden und Belgien

2.4.4 Zusammenfassung der Ergebnisse

E-learning in der Fluglotsenausbildung wurde im Vergleich zu Simulatortraining oder Klassenunterricht als weniger effektiv beurteilt. Aufgrund der nur einmalig bei der Entwicklung anfallenden Kosten konnte es jedoch ökonomisch effizient eingesetzt werden. Die gut gestalteten *E-learning* Module haben sich in beiden Akademien bereits amortisiert oder werden dies in Kürze erreichen. Die zum Zeitpunkt der Datenerhebung ermittelten ROIs für die schwedische Flugsicherungsakademie lagen zwischen 30 und 260% und für die belgische zwischen 23 und 120%, d.h. die ökonomische Bilanz ist überwiegend positiv.

Die befragten *Trainees*, *Trainer* und Akademiemanager machten aber auch deutlich, dass sie Verbesserungsbedarf sehen. Die *Trainer* und Akademiemanager kritisierten insbesondere, dass die von EUROCONTROL entwickelten Module zu wenig die lokalen Gegebenheiten in den Mitgliedsstaaten berücksichtigen. So müssen die in Belgien oder Schweden gültigen Flugsicherungsverfahren im Unterricht gesondert ver-

mittelt werden. Wünschenswert wären daher flexible Module, in die lokale Gegebenheiten von den *Trainern* eingearbeitet werden können. Auch sprachliche Unterstützungsfunktionen – etwa ganze Module in der Landessprache oder zumindest Wörterbücher und Glossars – wurden als Verbesserungsoption genannt. Diese Einschätzung spiegelte sich in den Antworten der *Trainees* wider: Die schwedischen *Trainees* bewerteten die Module aus dem eigenen Haus (in schwedisch und mit schwedischen Gegebenheiten) besser. Das EUROCONTROL Modul *Radar theory* wurde zwar im Mittel als potenter Ersatz für Simulator und Klassenunterricht gesehen, jedoch bestand hier große Uneinigkeit. Beide *Trainee*gruppen kritisierten auch die mangelnde individuelle Anpassungsfähigkeit der Module, z.B. an den individuellen Lernfortschritt. Die Motivation zum *E-learning* durch Zielklarheit, Einbettung in den Lehrplan, *Feedback* und eine gute Lernatmosphäre scheint ein zentraler Verbesserungsansatz für die *Trainer* zu sein. Die technische Gestaltung dagegen wurde relativ gut beurteilt, d.h. die Module wurden als gut instruiert, gut verständlich und leicht zu navigieren beschrieben. Dieser Anforderungswechsel ist allen Beteiligten wohl bewusst und IANS hat bereits vor einem Jahr den Aufmerksamkeitsfokus von den *Trainees* auf die *Trainer* verschoben. In verschiedenen *Workshops* für *Trainer* wurden und werden Aspekte der Implementierung und Integration von *E-learning* sowie das Arbeiten mit *Feedback* diskutiert und vermittelt.

Die schwedischen *Trainees* standen *E-learning* generell kritischer gegenüber. Dies war im Vorfeld der Untersuchung nicht erwartet worden, im Gegenteil: Wir gingen davon aus, dass *E-learning* in Schweden aufgrund der langjährigen Erfahrung besser entwickelt und akzeptiert ist. Es ist davon auszugehen, dass der Länderunterschied ein subjektives Phänomen ist und nicht etwa qualitativ schlechteres *E-learning* in Schweden nahe legt. Vielmehr sind die schwedischen *Trainees* in Bezug auf *E-learning* verwöhnt, da sie individuell angepasste und in ihrer Sprache programmierte Module kennen. Sie sind daher kritischer in ihren Urteilen als die belgischen Kollegen.

Einig sind sich beide *Trainee*gruppen darin, dass mehr Zeit auf *E-learning* verwendet werden sollte, als dies bisher der Fall ist. Während die Befragten im Durchschnitt 9-20% ihrer Lernzeit damit zubrachten, hielten sie 20-25% für in Zukunft angemessen.

2.4.5 Resumée zur Kurzevaluation

Gegenüber den anderen beiden Evaluationen ergaben sich zwei neue Schwierigkeiten, die in Anleitungen für zukünftige Studien Berücksichtigung finden sollten:

1. War die Evaluation der *E-learning* Module für die Beteiligten in der schwedischen und belgischen Akademie sicherlich interessant, jedoch nicht von hoher Priorität. Da Belgocontrol ausschließlich EUROCONTROL Module verwendet und dies auch nicht in Frage stellt, wurde ein Nutzen für die eigene Organisation durch die Teilnahme an der Studie kaum gesehen. Das *Commitment* zur Studie war entsprechend gering. Bei Luftfahrtverket ergab die Studie immerhin auch Rückmeldung über hausinterne Programme und wahrscheinlich wurden aus diesem Grund zügig Fragebogen zurückgesendet und Daten zu Kosten bereitgestellt.
2. Die Kostenermittlung war besonders schwierig, da sie teilweise eine Datenermittlung außerhalb der betrachteten Organisation erforderte. Dies war für alle EUROCONTROL Module der Fall. Erst kurz vor der Berichtlegung konnte eine obere

Abschätzung der Kosten für die Nutzung von EUROCONTROL Modulen ermittelt werden.

Das eigentliche Ziel dieser dritten Evaluation, mit wenig Aufwand verlässliche Managemententscheidungen bezüglich der Trainingsmaßnahme *E-learning* zu ermöglichen, wurde jedoch erreicht: *E-learning* hat sich – unter Zuhilfenahme überwiegend konservativer Schätzungen von Kosten und Nutzen – in den untersuchten Organisationen als rentabel erwiesen und sollte daher weiter ausgebaut werden.

Für den Arbeitsschritt „Ressourcenkontrolle“ im HPM konnte wertvolle Hinweise zur Methodik der Effizienzbewertung als Bestandteil der Wirtschaftlichkeitsanalyse gewonnen werden. Dazu gehört insbesondere die Kalkulation von Effizienzkenwerten für Qualifizierungsmaßnahmen.

3 Gesamtdiskussion

Mit den drei vorliegenden empirischen Studien konnte das HPM erfolgreich auf seine Anwendbarkeit und Nützlichkeit getestet werden. Die wesentlichen, zu prüfenden Inhalte des HPM (Kapitel 1.4) waren:

1. Die grundlegende Differenzierung der Wirtschaftlichkeitsbetrachtung nach den drei Ebenen „Organisation“, „Prozess“ und „Funktion“, die aufeinander aufbauen.
2. Die Modellierung des Evaluationsprozesses in Form von sieben Arbeitsschritten von der Analyse, über die Entscheidung/Planung bis zur Wirtschaftlichkeitsbewertung.
3. die Trennung der Bewertung einer Personalmaßnahme in eine Wirkungseinschätzung und eine Bedeutungseinschätzung

zu 1. Die umfassende, ökonomische Evaluation von CISM hat die Differenzierung der drei Ebenen als sinnvoll und für die Wirtschaftlichkeitsbetrachtung aufschlussreich bestätigt. Die Parameter und Zusammenhänge der drei Ebenen konnten in einem spezifischen Datenmodell für den Bereich der Flugsicherung und die Intervention CISM operationalisiert werden. Diese Form der Darstellung von Ursache-Wirkungs-Zusammenhängen im Rahmen der ökonomischen Evaluation kann als wegweisend beurteilt werden, da sie eine konkrete und anschauliche Landkarte für die ökonomische Analyse und Bewertung von Verhaltensmaßnahmen schafft. Die drei Ebenen können so als miteinander verbundene Regulationskreise verstanden werden, so dass auch vermeintlich „kleine“ Wirkungen in einem größeren Zusammenhang zu sehen sind.

Die Datenerhebung für diese drei Ebenen erfordert einen interdisziplinären Forschungsansatz und die Beteiligung verschiedener Zielgruppen im Unternehmen. Dies hat die CISM-Studie eindrücklich gezeigt. Zu berücksichtigen sind in der Regel sowohl Verhaltensdaten, die z.B. über eine Befragung der Zielgruppe oder deren Vorgesetzte erhoben werden können, als auch Leistungsdaten. Letztere können sowohl durch eine Befragung als auch durch das *Controlling* oder *Performance-Management* im Unternehmen gewonnen werden. Die CISM-Studie zeigte, wie eine ökonomische Evaluation auch ohne objektive Daten möglich ist, auch wenn die Einbeziehung objektiver Daten grundsätzlich zu empfehlen ist.

Es ist in dieser Studie auch deutlich geworden, dass für jede der drei Ebenen jeweils Teilmodelle notwendig sind, um die Wirkungszusammenhänge aufzeigen zu können. Dies ist insbesondere für die Prozessebene wichtig, die die Verbindungslinie zwischen den Verhaltenswirkungen und den strategischen Zielen bildet. Dabei ist eine prozessuale Betrachtung der Leistungserstellung in allen drei Studien und vermutlich in vielen Fällen in der Praxis nicht offensichtlich und die Prozessabläufe sind im Unternehmen nicht visualisiert wie dies aus typischen Prozessdarstellungen, z.B. aus der industriellen Fertigung, bekannt ist. Wenn die Leistungserstellung eher als isolierte Einzelarbeit wie im *Call Center* strukturiert ist oder die Abläufe wie in der Flugsicherung variieren können, greift die klassische Prozessbeschreibung als Abfolge von Tätigkeiten

nicht. Hier bedarf es zukünftig anderer Formen der Prozessmodellierung nicht nur für die ökonomische Evaluation.

Die CISM-Studie hat sich in der Prozessbetrachtung ausschließlich auf die relevanten Prozessmerkmale wie Produktivität oder Sicherheit und die personelle Prozesssteuerung durch die Führungskraft bezogen. Ein sektorübergreifendes Prozessdesign, das die Auswirkungen von kritischen Ereignissen auf die Arbeitsabläufe hätte aufzeigen können, lag nicht vor. Für den Flugsicherungsprozess wäre aber beispielsweise bedeutsam, wie sich individuelle Verhaltensvariationen auf den sektorübergreifenden Arbeitsfluss auswirken und in welcher Form dies dann die Gesamtkapazität des Systems beeinflusst. Dann könnte auch erklärt werden, wann Veränderungen in der individuellen Informationsverarbeitung und der Qualität der Verkehrsplanung zu einer Reduzierung des gesamten Verkehrsflusses führen. Dies ist ein wichtiges, zukünftiges Forschungsfeld in der ökonomischen Evaluation. Das HPM könnte damit weiter vervollständigt werden.

Die beiden anderen, kleineren Studien konnten von der Differenzierung der drei Ebenen nicht profitieren. In beiden Fällen beschränkte sich die Betrachtung auf die Funktionsebene als klassisches, verhaltenswissenschaftliches Evaluationsfeld. Die Gründe dafür waren unterschiedlich. In der *Call Center* Studie gab es zum einen keine Dokumente und Ideen, wie die individuelle Tätigkeit als Prozess dargestellt werden könnte. Die Arbeit des Agenten wurde von den Verantwortlichen eher als Summe von verschiedenen Einzeltätigkeiten gesehen. Die Interventionen wurden außerdem als funktionsbezogene Qualifizierungsmaßnahmen eingesetzt. Eine Optimierung des Arbeitsablaufes war hier nicht das Ziel. Damit waren auch die abhängigen Variablen in der Evaluation ausschließlich funktionsorientiert. In der *E-learning* Studie lag der Grund für die fehlende Prozessbetrachtung darin, dass hier die Optimierung und Effizienzsteigerung des individuellen Lernvorgangs anvisiert wurde. Inwieweit dieses Lernen in den Arbeitsprozessen wirtschaftlich genutzt werden kann, lag außerhalb der Verantwortlichkeit der Ausbildungsakademien und war daher nicht Bestandteil der Fragestellung. Dies ist vielleicht charakteristisch für die Evaluation von Personalmaßnahmen in Organisationseinheiten, die nicht an der eigentlichen Wertschöpfung beteiligt sind bzw. die davon zeitlich und räumlich distanziert sind.

In der *E-learning* Studie hat sich eine weitere wichtige Betrachtungsebene der Evaluation ergeben, die ergänzend zu den drei HPM-Ebenen zu sehen ist: die Bewertung der Anwendungseffizienz einer Personalmaßnahme. Es war ein wichtiger Teil der Auswertung, inwieweit *E-learning* einen Lernvorgang wirtschaftlicher werden lässt, als andere Lernformen, und zwar unabhängig vom (späteren) Anwendungsnutzen im Feld. Diese vierte Ebene der Wirtschaftlichkeitsbetrachtung sollte in das HPM bereits in der Analysephase aufgenommen werden.

zu 2. Der Evaluationsprozess war in allen drei Studien gemessen am HPM-Prozessmodell unvollständig. Das HPM sieht folgende Prozessschritte vor:

- I. Analyse auf allen drei Ebenen beginnend mit der Organisationsebene (Schritt 1-3)

- II. Entscheidung für bestimmte Personalmaßnahmen und Planung sowie Umsetzung dieser Maßnahmen (Schritt 4)
- III. *Monitoring* der anfallenden Kosten während der Umsetzung (Schritt 5)
- IV. Analyse der Wirkung der Maßnahmen (Schritt 6)
- V. Analyse der Bedeutung der Maßnahmen (Schritt 7)

Es zeigte sich, dass in der Praxis eine Evaluation oft im Hinblick auf bereits durchgeführte oder zumindest geplante Maßnahmen erfolgt. Eine interventionsunabhängige Analyse, wie sie das HPM in den drei ersten Schritten vorsieht, erscheint eher unrealistisch. Die Evaluation beschränkte sich vielmehr in allen drei Fällen auf die Wirkungs- und Bedeutungsanalyse, wobei die drei in der Analysephase unterschiedenen Ebenen auch für diesen Evaluationsprozess von Bedeutung sind. Sie bilden damit die Struktur der Evaluation unabhängig vom konkreten Prozess und vom einzelnen Arbeitsschritt. Sowohl eine klare Trennung von Struktur und Prozess, als auch eine Fokussierung der Analyse auf die konkret geplanten Maßnahmen von Anfang an im Prozess erscheinen daher angemessen. Dies war bisher im HPM nicht vorgesehen.

Außerdem zeigen die Erfahrungen, dass es sinnvoll ist, unterschiedliche Zeitpunkte und damit Zielsetzungen einer ökonomischen Evaluation stärker im Modell zu verankern und diese unabhängig von der Prozessgestaltung und der Struktur zu betrachten. Bei den drei Studien handelte es sich um begleitende Evaluationen. Die Maßnahmen waren bereits eingesetzt, sollten aber durch die Evaluation entweder bestätigt oder optimiert werden. Neben der begleitenden Evaluation ist auch eine Wirtschaftlichkeitsbetrachtung vor der Durchführung einer Maßnahme denkbar und sinnvoll, z.B. bei der Entscheidung über verschiedene alternative Maßnahmen. Wirkungen müssten hier unter Verwendung des HPM-Strukturmodells prognostiziert werden. Die ökonomische Evaluation hätte Auswirkungen auf die Entscheidung, Planung und Umsetzung. Die Analyse auf den Ebenen „Organisation“, „Prozess“ und „Funktion“ wäre allerdings auf bereits in Betracht gezogene, alternative Interventionen zu beziehen und nicht, wie bisher im HPM vorgesehen, interventionsunabhängig. Schließlich ist auch gut vorstellbar, dass eine ökonomische Evaluation zum Abschluss einer Maßnahme durchgeführt wird, um die Wirkung einer Personalmaßnahme im Sinne einer „Bilanz“ zu dokumentieren.

- zu 3. Die Trennung in eine Wirkungs- und eine Bedeutungsanalyse war vor allem in der CISM-Studie hilfreich und relevant. So konnte zwischen den Verhaltenswirkungen von CISM und den Konsequenzen dieser Verhaltensauswirkungen für das Unternehmen unterschieden werden. Allerdings lassen sich auch hier aus den Erfahrungen Überlegungen zu einer Optimierung ableiten. Die Unterscheidung dieser beiden Bewertungsstufen ist nur eine umgangssprachliche Vereinfachung des 3-Ebenen-Strukturmodells (Organisation, Prozess, Funktion), wobei die Wirkungsanalyse sich auf die Funktionsebene bezieht und die Bedeutungsanalyse auf die Prozess- und Organisationsebene. Die ökonomische Bewertung individueller Verhaltensänderungen über Schätzverfahren, wie es in der CISM-Studie mit der *Performance*-Kurve erprobt wurde, bietet zwar auch Anhaltspunkte für eine Bedeutungsanalyse. Diese sollte aber durch konkrete Wirkungsmodelle, die alle Ebenen einbeziehen, abgesichert und über-

prüft werden. Insofern ist eine Verwendung des Strukturmodells auch für die Wirtschaftlichkeitsbewertung zu empfehlen. Dabei muss auch hier einschränkend gesagt werden, dass in den beiden kleineren Studien eine Wirtschaftlichkeitsbewertung unter Einbeziehung von Prozessmerkmalen und strategischen Zielen aufgrund der oben angeführten Bedingungen nicht umgesetzt werden konnte.

Die *E-learning* Studie zeigte, dass als vierte Ebene der Wirtschaftlichkeitsbewertung eine Effizienzbetrachtung der Maßnahme selbst in Frage kommt. Diese könnte die reine Ressourcenkontrolle, wie sie im HPM vorgesehen ist, ersetzen.

Alle drei Studien haben zudem einen Weg aufgezeigt, wie eine ökonomische Evaluation auch ohne Kontrollgruppe und ohne *Pretests* möglich ist. Dieser Mangel konnte in großen Teilen kompensiert werden, indem verschiedene Datenquellen im Sinne der *concurrent validity* die Gültigkeit des Gesamtbildes bestätigten. Weitere Behelfslösungen waren die retrospektiv erhobenen Daten in der CISM-Studie, die quasi als *Pretest* fungierten. Auch wenn dies nicht dem wissenschaftlichen Goldstandard entspricht, hat sich diese Vorgehensweise bewährt und kann als aussichtsreicher Kompromiss in dem Konflikt zwischen den Ansprüchen der Wissenschaft und den Möglichkeiten der Praxis bei der Evaluation von HF/HR/T-Maßnahmen angesehen werden.

Die Erfahrungen aus den drei Studien und die angeführten Überlegungen bieten Ansatzpunkte für eine Überarbeitung des HPM. Ein überarbeitetes HPM sollte drei Aspekte berücksichtigen:

1. Die Nutzung der drei Analyseebenen sowie einer weiteren „Effizienzebene“ als grundsätzliches Strukturmodell.
2. Die Modellierung des Evaluationsprozesses mit engem Bezug zu den beabsichtigten oder realisierten Maßnahmen.
3. Die Konkretisierung von Struktur und Prozess in Abhängigkeit vom Zeitpunkt und der Zielsetzung der Evaluation.

Das Strukturmodell des überarbeiteten HPM

Die vorliegenden Studien haben die unter 1. genannten drei Wirkungsebenen bestätigt. Die *E-learning*-Evaluation zeigte die Notwendigkeit einer wirkungsunabhängigen Effizienzebene auf, die sich nur auf die Kosteneffizienz der Maßnahme selbst bezieht, ohne den Nutzen im Wirkungsfeld zu betrachten. Darüber hinaus rückt eine weitere Wirkungsebene in den letzten 1-2 Jahren immer mehr in den Blickpunkt des *Personalmanagements*, die bereits in der Literaturübersicht zu diesem Projekt und im Kapitel 1.3.1 dieses Forschungsberichtes ihren Platz hatte, im HPM aber noch nicht berücksichtigt wurde, da sie vermeintlich nur für externe Anspruchsgruppen relevant ist: das Humankapital. Es zeigt sich in der aktuellen Entwicklung immer mehr, dass die Steigerung des Humankapitals als langfristiger Kompetenzvorteil des Unternehmens eine wichtige Zielgröße für das *Personalmanagement* werden wird, an der sich

auch Personalmaßnahmen messen lassen müssen. Das Humankapital ist in den drei Wirkungsebenen des HPM nur insofern enthalten, als es sich in strategischen Zielgrößen, Prozesskennzahlen oder individuellen Kompetenzmerkmalen widerspiegelt.

Das überarbeitete Strukturmodell sieht daher fünf Wirtschaftlichkeitsebenen vor. Eine ökonomische Evaluation kann, muss aber nicht alle fünf Ebenen berücksichtigen:

| Ebene | Bewertungskriterium für die Evaluation |
|--------------|---|
| Humankapital | Steigerung der nachhaltigen Kompetenzvorteile des Unternehmens (Kompetenzvorteile, Engagement, Bindung, Attraktivität des Unternehmens) |
| Organisation | Beitrag zur Strategieumsetzung (Deckungsbeiträge, Personalkosten, Kundenzufriedenheit etc.) |
| Prozess | Beitrag zur Prozessoptimierung (Prozesskosten, Durchlaufzeiten, Prozessqualität, Produktivität) |
| Funktion | Beitrag zur individuellen Leistungsoptimierung (Kompetenzen, Verhalten, Leistung der Zielgruppe) |
| Maßnahme | Beitrag zur Verbesserung der Effizienz der Maßnahme (Kosten pro Lerneinheit, Kosten pro Einstellung etc.) |

Abbildung 62: Strukturmodell zur ökonomischen Evaluation

Die fünf Ebenen sind wechselseitig miteinander verbunden. In der Analyse und Bewertung kann ein *Top-down*-Vorgehen (von der Strategie ausgehend) mit einem *Bottom-up*-Vorgehen (von der einzelnen Maßnahme ausgehend) kombiniert werden, um die Ursache-Wirkungs-Zusammenhänge zu erschließen.

Prozessmodell der ökonomischen Evaluation

Unabhängig von der Fragestellung und Zielsetzung der Evaluation erscheint aus den Erfahrungen der Untersuchungen heraus folgendes grundlegende Vorgehen in der ökonomischen Evaluation zielführend zu sein:

1. Modellierung der Parameter, der Zusammenhänge und der Erfolgsfaktoren

Analog zur CISM-Studie sollte der Evaluationsprozess mit einer detaillierten Analyse der einzubeziehenden Faktoren und deren möglichen Wirkungszusammenhängen beginnen. Das HPM-Strukturmodell bietet hier eine Landkarte als Startpunkt und Orientierungshilfe für die Modellierung. An diesem Ausgangspunkt der Evaluation sind zunächst die zu evaluierenden Maßnahmen genau zu beschreiben und ihre möglichen Einflüsse auf das gesamte Leistungssystem, d.h. auf die verschiedenen Wirkungsebenen, zu durchdenken. Insbesondere sollten hier die Erfolgsfaktoren und die Engpassfaktoren im Leistungssystem

herausgearbeitet werden. Die Modellierung kann z.B. auf der Basis von Experteninterviews in der Organisation unter Nutzung vorhandener Dokumente, die die relevanten Maßnahmen beschreiben (z.B. *Trainingsleitfäden*), durchgeführt werden. Hier hat sich der interdisziplinäre Ansatz in der CISM-Studie bewährt, bei dem Experten verschiedener Fachrichtungen (Personal, *Controlling*, Führungskräfte aus der Linie) einbezogen wurden.

2. Planung und Konzeption der Datenerhebung für die Evaluation

Darauf aufbauend ist ein Untersuchungsdesign zu erstellen, mit dem die Wirkung der Maßnahme zuverlässig und valide überprüft werden kann. Die einzubeziehenden Faktoren sind zu operationalisieren und die Messinstrumente müssen konzipiert werden. Das erarbeitete Strukturmodell dient wie bei der CISM-Studie in dieser Phase als *Masterplan*. Die in den drei Studien entwickelten Interviewleitfäden und Fragebogen, inklusive der *Performance*-Kurve der CISM-Studie, sind als Beispiele im Anhang dieses Berichts enthalten.

3. Erhebung und Auswertung der Daten

Die Datenerhebung erfolgt unter Berücksichtigung der geltenden Datenschutzbestimmungen und in Abstimmung mit der organisierten Mitarbeiterbeteiligung (Personal- oder Betriebsrat). Entsprechend dem entwickelten Untersuchungsdesign werden nun die Messinstrumente eingesetzt, ggf. unter Einbeziehung einer Kontrollgruppe und/oder einem *Pretest*, oder auch als Längsschnitt, d.h. durch Mehrfachmessung im Verlauf einer Maßnahme. Wichtig ist es, alle Daten so kompatibel zu machen, dass in der Auswertung eine Analyse über die verschiedenen Wirkungsebenen möglich ist. In der CISM-Studie wurden beispielsweise die Funktionsträger (Fluglotsen) auch nach Prozessmerkmalen gefragt, um Zusammenhänge zwischen der funktionalen Ebenen und der Prozessebene herzustellen. Die Auswertung sollte nicht nur Ergebniskennzahlen enthalten, sondern das Strukturmodell im Rahmen einer Pfadanalyse oder einer Korrelationsanalyse wieder aufgreifen.

4. Rückkopplung der Daten und Ergebnisinterpretation mit den Experten

Ein wichtiger Schritt für eine wirkungsvolle Evaluation ist die abschließende Rückkopplung der Befunde an die Experten und Führungskräfte und die gemeinsame Interpretation der Daten. Auch dies war ein Erfolgsfaktor in der CISM-Hauptstudie, da so die verschiedenen Fachrichtungen im Unternehmen die Möglichkeit hatten, den Ergebnissen selbst eine Bedeutung zu geben und die relevanten Zusammenhänge im Detail nachzuvollziehen.

Das erweiterte HPM-Strukturmodell und dessen Operationalisierung für den konkreten Anwendungsfall ist die Grundlage für alle vier Schritte der Evaluation. Welche der fünf Ebenen (Humankapital, Organisation, Prozess, Funktion, Maßnahme) in einer Evaluationsstudie relevant sind bzw. die höhere Priorität haben, hängt von der Zielsetzung der Evaluation ab und davon, welche Verantwortungsbereiche sich für die Ergebnisse der Evaluation interessieren. In der *E-learning* Studie gab es in erster Linie ein hohes Interesse an der Kosteneffizienz der neuen Maßnahmen. Die vier Wirkungsebenen interessierten weniger. In der CISM-Studie war dagegen die Kosteneffizienz nicht relevant, weil keine Alternative für das CISM-Programm gesehen wurde. In der *Call Center* Studie gab es keine Kostendaten zu den Programmmodulen, daher war eine Effizienzanalyse nicht möglich. In allen drei Studien wurde die Steige-

rung des Humankapitals nicht untersucht. Die humankapitalorientierte Personalarbeit steckt noch in den „Kinderschuhen“ und nicht jedes Unternehmen wird sich daran orientieren, das eigene Humankapital strategisch auszubauen. Daher wird diese Nutzenperspektive in der ökonomischen Evaluation in naher Zukunft vermutlich noch nicht häufig aufgegriffen werden.

Zeitpunkt der ökonomischen Evaluation

Als Zeitpunkt der Evaluation sind grundsätzlich drei Varianten denkbar. Die Evaluation vor der Entscheidung für eine Maßnahme, die begleitende Evaluation während der Anwendungszeit, um die Maßnahme fortlaufend zu optimieren und ihre Zweckmäßigkeit zu überprüfen, sowie die nachträgliche Kosten-Nutzenbewertung zum Zweck der Dokumentation. Je nach Zeitpunkt ergeben sich verschiedene Anwendungsdetails im Evaluationsprozess und der Struktur.

1. Evaluation vor der Entscheidung (prognostische Evaluation)

Die ökonomische Evaluation vor der Entscheidung für eine oder mehrere Personalmaßnahmen erfordert eine Prognose von Kosten und Nutzeneffekten. Dies kann z.B. in Form einer Pilotstudie erfolgen, bei der die Wirkung an einer ausgewählten Stichprobe oder im Rahmen von *Benchmarking*analysen untersucht wird, bei denen die Wirkungen in anderen Organisationen herangezogen werden. Möglich ist auch die Simulation von Wirkungen im „Labor“ oder die Anwendung von Szenariotechniken, bei der die Sensibilität wirtschaftlicher Kennzahlen im Unternehmen für Verhaltensänderungen bei der ausgewählten Zielgruppe getestet wird. Dies wäre auch eine interessante Evaluationsmethode für CISM bei Fluglotsen. Der Evaluationsprozess entspricht in diesem Fall den vier genannten Schritten, wobei in der Modellierungsphase die verschiedenen, alternativen Maßnahmen und ihre möglichen Einflüsse auf das Leistungssystem skizziert werden müssen. Die Datenerhebung in einer Pilotstudie oder Machbarkeitsstudie ist in der Regel nicht so umfangreich und detailliert wie in einer begleitenden

Evaluation. Andererseits sollte gerade diese Form der Evaluation mit besonderer Sorgfalt und unter Berücksichtigung aller Wirkungsstufen erfolgen, da die prognostische Evaluation schwerwiegende Investitionsfehler vermeiden hilft, während in der begleitenden Evaluation i.d.R. nur kleine Anpassungsmaßnahmen gefragt sind, um die Effizienz einer bereits auf den Weg gebrachten Personalmaßnahme zu steigern. Die Detailplanung und Umsetzung der Personalmaßnahmen erfolgt hier nach Schritt 4 des Evaluationsprozesses. An eine prognostische Evaluation kann sich eine begleitende Evaluation anschließen.

2. Evaluation während der Umsetzung (begleitende Evaluation)

Eine begleitende Evaluation ist vermutlich die am weitesten verbreitete Form der Wirtschaftlichkeitsanalyse. Auch die drei in diesem Bericht dokumentierten Studien gehören in diese Kategorie. Die Zielsetzung ist in der Regel die Optimierung der laufenden Maßnahme oder auch, wie z.B. bei CISM, die Dokumentation der Wirtschaftlichkeit für das *Management* und für andere potenzielle Nutzer von CISM in anderen Flugsicherungsorganisationen und anderen Branchen. Wie die Studien zeigen, hängt die Verwendung des Strukturmodells von der Zielsetzung der Verantwortungsträger als Adressaten der Evaluation ab. Das Vorgehen entspricht dem aufgezeigten Prozessmodell. Je vollständiger das Strukturmodell verwandt wurde, desto größer ist auch die Bedeutung der Prozessmodellierung

und der Dateninterpretation mit den Experten und Führungskräften. Es zeigte sich, dass sich an diesen beiden Schritten der Wirkungsgrad der Evaluation festmachen lässt. Die intensive Modellierung, die Einbeziehung der Experten in den Prozess und der Austausch mit den Führungskräften über die Bedeutung der Daten waren die eigentlichen Erfolgsfaktoren in der CISM-Studie.

3. Evaluation nach Abschluss der Maßnahme (dokumentierende Evaluation)

Die Evaluation nach Beendigung einer Personalmaßnahme ist dann sinnvoll, wenn die Wirtschaftlichkeit einer umfassenden Investition für das *Management* dokumentiert, die Wirksamkeit einer Maßnahme wissenschaftlich nachgewiesen und der Fachwelt präsentiert oder Implikationen für zukünftige Anschlussmaßnahmen aufgezeigt werden sollen. Abhängig von der Zielsetzung und dem Adressatenkreis sind die verschiedenen Ebenen des Strukturmodells relevant und die vier Arbeitsschritte auszugestalten.

Die Integration der oben beschriebenen Aspekte in einem neuen Gesamtmodell zeigt Abbildung 63.

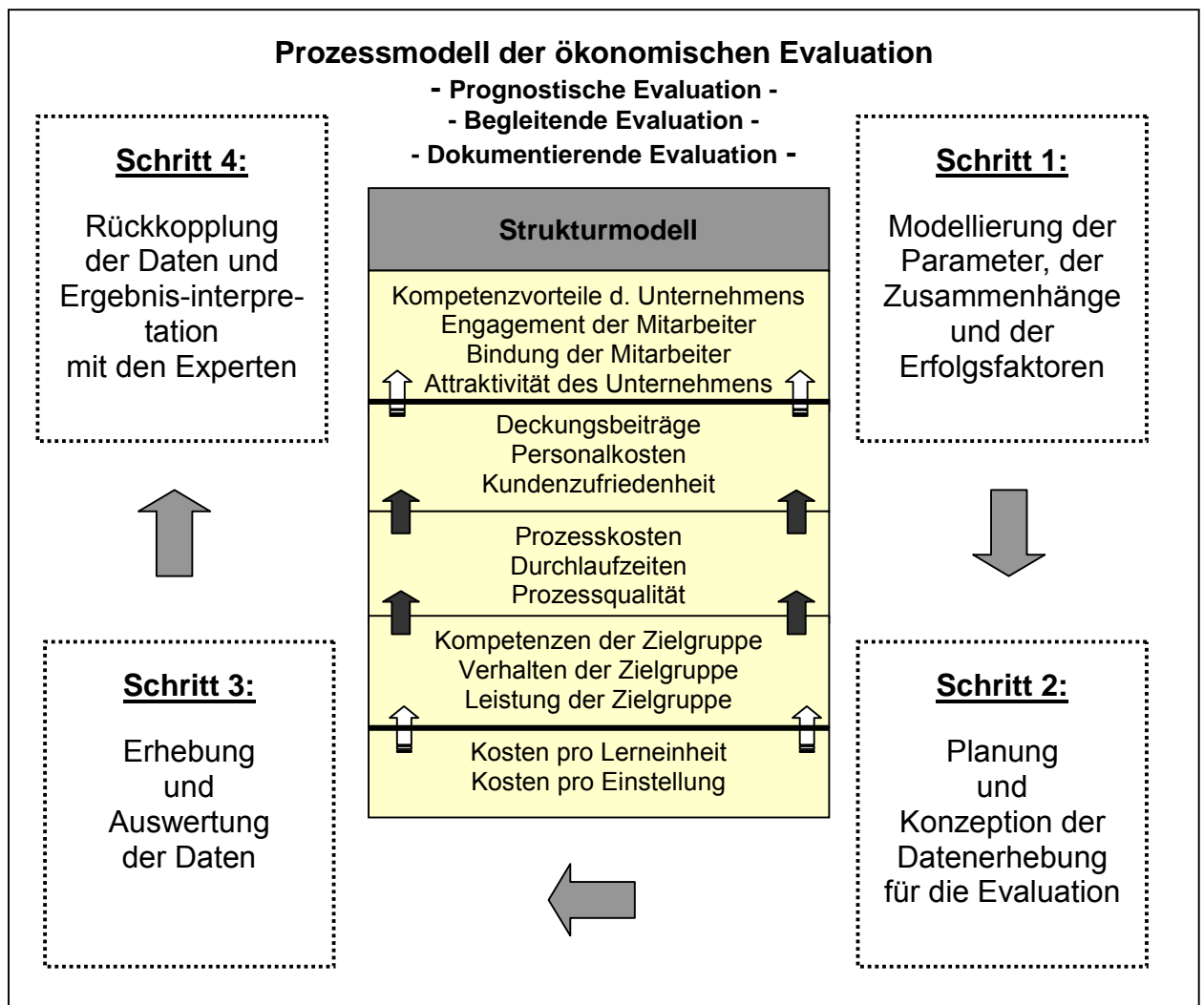


Abbildung 63: Das erweiterte HPM als differenziertes Prozess- und Strukturmodell

Die Erkenntnisse der drei Studien und das erweiterte HPM sollten konsequent weiter beforscht werden. Ein Schwerpunkt ist dabei in der Modellierung der Prozessebene und den Wirkungszusammenhängen zwischen funktionsbezogener Verhaltensvariabilität und Prozessstabilität zu setzen.

Mit den Produkten und Dokumenten aus diesem Forschungsprojekt (s. Kapitel 1.3) hoffen wir, Impulse für die ökonomische Evaluation von HF/HR/T-Maßnahmen, die Evaluationsforschung im Allgemeinen und die praxisorientierte Anwendung und Verfeinerung der entwickelten Modelle und Methoden zu geben.

4 Literaturverzeichnis

- [1] Boudreau, J.W. & Ramstad, P.M. (2003). Strategic industrial and organisational psychology and the role of utility analysis models (pp. 193-221). In W.C. Borman, D.R. Ilgen & R.J. Klimoski (Eds.), *Handbook of Psychology, Vol. 12, Industrial and Organizational Psychology*. New York: John Wiley and Sons.
- [2] Brähler, E. & Scheer, J. W. (1983). *Der Gießener Beschwerdefragebogen (GGB)*. Stuttgart, Wien: Hans Huber.
- [3] Deuchert, I. & Eißfeld, H. (1998). Potenzialanalyse in der Flugsicherung. In M. Kleinmann & B. Strauße (Hrsg.), *Potenzialfeststellung und Personalentwicklung*. Göttingen: Verlag für Angewandte Psychologie.
- [4] EUROCONTROL (2004a). *An Assessment of Air Traffic Management in Europe during the Calendar Year 2003*. Brussels: EUROCONTROL Performance Review Report 7.
- [5] EUROCONTROL (2004b). *Evaluation of human factors and resource programmes with economic evaluation methods*. Contributors: B. Köper, S. Pennig & J. Vogt. Brussels: EUROCONTROL.
- [6] EUROCONTROL (2005). *The First European Air Traffic Controller Selection Test FEAST*. <http://www.eurocontrol.int/humanfactors/feast.html>
- [7] Fleishman, E. A. & Reilly, M. E. (1992). *Handbook of human abilities—definitions, measurements and job task requirements*. Palo Alto, CA, USA: Consulting Psychologists Press.
- [8] Goetzel, R.Z., Juday, T.R., Ozminkowski, R.J. (1999). *What's the ROI? A systematic review of Return on Investment studies of corporate health and productivity management of worksite health*.
- [9] Kaplan, R.S. & Norton, D.P. (1996). *The Balanced Scorecard – translating strategy into action*. Boston: Harvard Business School Press.
- [10] Kastner, M. (2006). *Vertrauens-Fehler-Lern-Innovations-Gesundheits-Kultur*. Vortrag auf der Frühjahrstagung der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin Dortmund.
- [11] Kirkpatrick, D. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco: Barrett-Koehler.
- [12] Leonhardt, J. & Vogt, J. (2006). Critical Incident, Critical Incident stress, post traumatic stress disorder – definitions and underlying neurobiological processes (pp 43-52). In J. Leonhardt & J. Vogt (Editors) *Critical Incident Stress Management CISM in aviation. Set heading to the positive*. Aldershot: Ashgate.
- [13] Pennig, S. & Vogt, J. (2006). *Entwicklung einer Evaluationsroutine zur Prüfung der Nachhaltigkeit von Vorhaben im Rahmen des Modellprogramms zur Bekämpfung arbeitsbedingter Erkrankungen*. BAuA Forschungsprojekt F2145. Zwischenbericht zum 1.9.2006.

- [14] Pennig, S., Kremeskötter, N., Nolle, T., Koch, A., Maziul, M. & Vogt, J. (2006). Ökonomische Evaluation von Personalressourcen und Personalarbeit. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin FB 1070. Bremerhaven: Wirtschaftsverlag NW.
- [15] Phillips, J.J. (1996). Measuring ROI: The 5th level of evaluation. Technical and Skills Training, April, 1996.
- [16] Reason, J. (1986). Managing the risks of organizational accidents. Aldershot: Ashgate
- [17] Sars, M.B. (2006). Bridging the Rubicon of Evaluating Human Resources Connecting Human Resource Management to Strategic Steering through Programme Evaluation Linking "Intangible" Performance Effects with Organisational Effectiveness. Unpublished Master Thesis in Psychology, University of Copenhagen.
- [18] Turner, C. & Naismith, A. (1989). The emotional response to stress test (ERST). Psychology Dept. Report, University of Glasgow.
- [19] Vogt, J., Adolph, L., Ayan, T., Udovic, A. & Kastner, M. (2002). Stress in modern air traffic control systems and potential influences on memory performance. *Journal of Human Factors and Aerospace Safety*, 2(4), 355-378.
- [20] Vogt, J., Hagemann, T. & Kastner, M. (2006). The impact of workload on heart rate and blood pressure in en-route and tower air traffic control. *The Journal of Psychophysiology*, 20(4), 297-314.
- [21] Vogt, J., Leonhardt, J. Köper, B. & Pennig, S. (2004). Economic evaluation of the Critical Incident Stress Management Program. *The International Journal of Emergency Mental Health* 6(4), 185-196.
- [22] Vogt, J., Maziul, M., Sars, M. & Petter, B. (in press). Wissens- und Kulturaustausch zwischen Dänemark und Deutschland. In M. Kastner, E. Neumann-Held & C. Reick (Hrsg.) *Kultursynergien oder Kulturkonflikte*. Lengerich: Pabst Science Publishers.

5 Abbildungs- und Tabellenverzeichnis

| | Seite |
|--|-------|
| 5.1 Abbildungen | |
| <u>Abbildung 1:</u> <u>Das Human Resources Performance Modell</u> | 19 |
| <u>Abbildung 2:</u> <u>Einbindung des Verhaltensmanagements in das Prozessmanagement</u> | 24 |
| <u>Abbildung 3:</u> <u>Evaluationsmodell und Fragenbereiche der Machbarkeitsstudie (die Zahlen verweisen auf die Fragenummer im Fragebogen bzw. Interview)</u> | 30 |
| <u>Abbildung 4:</u> <u>Die verschiedenen Auskunftgeber und Informationsbereiche in der Evaluation des Critical Incident Stress Managements der Flugsicherung</u> | 33 |
| <u>Abbildung 5:</u> <u>Die Benchmark-Evaluation verbindet die Individual- mit der Organisationsebene und überschreitet so den psychoökonomischen Rubikon (SV: Supervisor, COS: Chief of Section)</u> | 34 |
| <u>Abbildung 6:</u> <u>Steuerungskreisläufe im Flugsicherungsprozess (DB: Deckungsbeitrag, EBG: Einsatzberechtigungsgruppe, NL: Niederlassung, UZ: Unternehmenszentrale)</u> | 37 |
| <u>Abbildung 7:</u> <u>Fragebogenrücklauf nach Niederlassungen (CI: Critical Incident, FB: Fragebogenabschnitt)</u> | 49 |
| <u>Abbildung 8:</u> <u>Erhaltene Peer-Fragebogen nach Niederlassungen</u> | 50 |
| <u>Abbildung 9:</u> <u>Prozentanteil der Fluglotsen mit geringer (0-25% und 26-50%), moderater (51-75%) und starker (76-100%) Beunruhigung/Aufregung durch ihre kritischen Ereignisse</u> | 52 |
| <u>Abbildung 10:</u> <u>Von Fluglotsen angegebener Belastungszeitraum nach dem kritischen Ereignis</u> | 53 |
| <u>Abbildung 11:</u> <u>Peer-Einschätzung der Belastungsdauer nach dem kritischen Ereignis</u> | 53 |
| <u>Abbildung 12:</u> <u>Art und Ausmaß der verhaltensbezogenen, emotionalen und körperlichen Beeinträchtigung (0 nicht, 1 kaum, 2 einigermaßen, 3 erheblich, 4 stark aufgetretene Beschwerden)</u> | 54 |
| <u>Abbildung 13:</u> <u>Von den Peers berichtete Beeinträchtigungen der Fluglotsen (0 nicht, 1 kaum, 2 einigermaßen, 3 erheblich, 4 stark aufgetreten)</u> | 55 |

| | | |
|----------------------|---|----|
| <u>Abbildung 14:</u> | <u><i>Performance</i>-Kurve (Bewältigung Arbeitslast/Arbeitslast) der Fluglotsen vor und nach dem kritischen Ereignis (CI; <i>Critical Incident</i>)</u> | 57 |
| <u>Abbildung 15:</u> | <u>Leistung der durch <i>Critical Incidents</i> betroffenen Fluglotsen in Prozent (<i>Performance</i>-Kurve) aus Sicht der <i>Peers</i> in den ersten Tagen nach dem Ereignis (Leistung 0 = Tag des <i>Critical Incident</i>; Leistung 1= 1 Tag nach CI; Leistung 2= 2 Tage nach CI Leistung 3 5= 3-5 Tage nach CI; Leistung 6 14= 6-14 Tage nach CI)</u> | 58 |
| <u>Abbildung 16:</u> | <u>Bewältigungsmechanismen der Fluglotsen nach dem kritischen Ereignis</u> | 59 |
| <u>Abbildung 17:</u> | <u>Nutzung von CISM in den verschiedenen Niederlassungen der DFS</u> | 60 |
| <u>Abbildung 18:</u> | <u>Ausmaß der Unterstützung der generellen Bewältigungsmechanismen durch CISM in Prozent</u> | 61 |
| <u>Abbildung 19:</u> | <u>Unterstützung der generellen Bewältigungsmechanismen aus Sicht der <i>Peers</i> in Prozent</u> | 62 |
| <u>Abbildung 20:</u> | <u>Beitrag der verschiedenen Einflussfaktoren zur Erholung in Prozent</u> | 63 |
| <u>Abbildung 21:</u> | <u>Veränderung der Fähigkeiten nach einem kritischen Ereignis in Prozent</u> | 64 |
| <u>Abbildung 22:</u> | <u>Veränderung der Tätigkeiten nach einem kritischen Ereignis in Prozent</u> | 65 |
| <u>Abbildung 23:</u> | <u>Auswirkungen der Beeinträchtigung nach kritischen Ereignissen auf den Flugsicherungsprozess</u> | 67 |
| <u>Abbildung 24:</u> | <u>Die verschiedenen Interventionsformen der Prozesssteuerung nach einem kritischen Ereignis</u> | 69 |
| <u>Abbildung 25:</u> | <u><i>Performance</i>-Kurven nach Interventionsformen (CI: <i>Critical Incident</i>)</u> | 70 |
| <u>Abbildung 26:</u> | <u>Tätigkeitseinschränkungen in den sechs Interventionsgruppen</u> | 71 |
| <u>Abbildung 27:</u> | <u>Ergebniswirkungen der verschiedenen Interventionsformen</u> | 72 |
| <u>Abbildung 28:</u> | <u><i>Performance</i>-Kurve für die CISM-Gruppen (CI: <i>Critical Incident</i>)</u> | 74 |
| <u>Abbildung 29:</u> | <u>Häufigkeit von <i>déjà-vus</i> bei Fluglotsen mit mehr als zwei Jahre zurückliegendem kritischen Ereignis</u> | 78 |

| | | |
|----------------------|--|----|
| <u>Abbildung 30:</u> | <u>Beeinträchtigungen durch <i>déjà-vu</i>-Situationen (1 nicht, 2 kaum, 3 einigermaßen, 4 erheblich, 5 stark aufgetretene Beeinträchtigungen)</u> | 79 |
| <u>Abbildung 31:</u> | <u>Kulturprofil der DFS anhand von Adjektiven (1: trifft genau zu, 6: trifft nicht zu)</u> | 81 |
| <u>Abbildung 32:</u> | <u>Veränderung des Kulturprofils der DFS durch CISM (+2 stark verbessert, -2 stark verschlechtert)</u> | 82 |
| <u>Abbildung 33:</u> | <u>Beitrag verschiedener HF/HR/T-Programme zur Veränderung der Sicherheitskultur der DFS CBT <i>Computer Based Training</i> für Flugsicherungsverfahren seit 1999</u> | 83 |
| <u>Abbildung 34:</u> | <u>Wirkungskette der Reaktionen auf ein kritisches Ereignis (mit CISM links in grün und ohne CISM rechts in rot)</u> | 84 |
| <u>Abbildung 35:</u> | <u>Datenmodell für die Programmevaluation</u> | 88 |
| <u>Abbildung 36:</u> | <u>Pfadmodell zur Programmevaluation im untersuchten <i>Call Center</i></u> | 89 |
| <u>Abbildung 37:</u> | <u>Prozentuale Anzahl von erfolgskritischen Items je Anforderungsbereich (MA: Mitarbeiter, FK: Führungskräfte)</u> | 92 |
| <u>Abbildung 38:</u> | <u>Erfolgskritische Anforderungen im <i>Call Center</i> dargestellt im Pfadmodell (AIS: Allgemeines Informationssystem)</u> | 93 |
| <u>Abbildung 39:</u> | <u>Kompetenzwerte für die drei Qualifikationsbereiche in der Selbst- und Fremdwahrnehmung (MA: Mitarbeiter, FK: Führungskräfte)</u> | 94 |
| <u>Abbildung 40:</u> | <u>Hohe Kompetenzwerte im <i>Call Center</i> nach dem PE-Programm (blau: in Führungskräfte­sicht über 75%, rot: in Mitarbeitersicht über 80%, schwarz: in beiden Sichten hoch ausgeprägt)</u> | 95 |
| <u>Abbildung 41:</u> | <u>Niedrige Kompetenzwerte im <i>Call Center</i> nach dem PE-Programm (blau: in Führungskräfte­sicht unter 50%, schwarz: in beiden Sichten unter 60%, unterstrichen: erfolgskritisch in Führungskräfte­sicht)</u> | 96 |
| <u>Abbildung 42:</u> | <u>Zielbeiträge der Programmmodule zu den Kompetenzwerten im <i>Call Center</i> (blau: Mittelwert der Führungskräfte, rot: Mittelwert der Mitarbeiter)</u> | 97 |
| <u>Abbildung 43:</u> | <u>Anforderungsmerkmale mit hohen Zielbeiträgen der Programmmodule aus Mitarbeitersicht (grün: Fachanweisung; rot: Einarbeitung; blau: Kommunikationsschulung; schwarz: Kommunikationsschulung + praktische Erfahrung)</u> | 98 |

| | | |
|----------------------|---|-----|
| <u>Abbildung 44:</u> | <u>Kompetenzen, deren Realisierung durch das PE-Programm von Mitarbeitern und Führungskräften mit über 30% Differenz beurteilt wird (grün: Fachanweisung; rot: Einarbeitung; blau: Kommunikationsschulung; schwarz: Kommunikationsschulung + praktische Erfahrung)</u> | 99 |
| <u>Abbildung 45:</u> | <u>Zielbeiträge des Entwicklungsprogramms (Werte links der Schrägstriche) bei niedrigen Kompetenzwerten (Werte rechts der Schrägstriche) in erfolgskritischen Anforderungsmerkmalen aus Führungssicht</u> | 100 |
| <u>Abbildung 46:</u> | <u>Qualität der Tätigkeitsausübung und der Ergebnisse aus Sicht der Führungskräfte (FK) und Mitarbeiter (MA)</u> | 102 |
| <u>Abbildung 47:</u> | <u>Zielbeiträge der Programmmodule zu Tätigkeits- und Ergebnisqualität aus Sicht der Führungskräfte (Werte links der Schrägstriche) und der Mitarbeiter (Werte rechts der Schrägstriche); die entsprechenden Werte der Kompetenzbereiche in hellgrau zeigen, dass die Befragten Zusammenhänge von Kompetenzen, Tätigkeits- und Ergebnisqualitäten realisieren</u> | 103 |
| <u>Abbildung 48:</u> | <u>Tätigkeits- und Ergebnismerkmalen, deren Realisierung durch das PE-Programm aus Sicht der Führungskräfte (Werte links der Schrägstriche) und der Mitarbeiter (Werte rechts der Schrägstriche) über 30% abweicht</u> | 104 |
| <u>Abbildung 49:</u> | <u>Einflüsse der Kontextfaktoren auf die Tätigkeitsausübung und die Ergebnisqualität (Zahlen als Querverweis zu den Kontextfaktoren; rot = negativer Einfluss)</u> | 105 |
| <u>Abbildung 50:</u> | <u>Korrelationen im Pfadmodell</u> | 108 |
| <u>Abbildung 51:</u> | <u>Mittelwerte selbst berichteter Fähigkeiten in Prozent für 13 schwedische und 14 belgische <i>Trainees</i> sowie 20 ausgebildeter Fluglotsen (ATCO)</u> | 116 |
| <u>Abbildung 52:</u> | <u>Mittlere selbst berichtete Qualität der Arbeitstätigkeiten in Prozent für 13 schwedische und 14 belgische <i>Trainees</i> sowie 20 ausgebildeter Fluglotsen (ATCO)</u> | 117 |
| <u>Abbildung 53:</u> | <u>Mittlere selbst berichtete Leistung in Prozent für 13 schwedische und 14 belgische <i>Trainees</i> sowie 20 ausgebildeter Fluglotsen (ATCO)</u> | 117 |
| <u>Abbildung 54:</u> | <u>Links: Mittlere Zeit in Prozent, die schwedische <i>Trainees</i> auf <i>E-learning</i>, Klassenunterricht und Simulatortraining verwendeten Rechts: Mittlerer Lerneffekt (Nutzen) in Prozent, den die schwedischen <i>Trainees</i> <i>E-learning</i>, Klassenunterricht und Simulatortraining zuwiesen</u> | 118 |

| | | |
|----------------------|---|-----|
| <u>Abbildung 55:</u> | <u>Mittlerer Quotient und Standardfehler von wahrgenommenem Lerneffekt pro eingesetzte Zeit in der schwedischen Stichprobe</u> | 119 |
| <u>Abbildung 56:</u> | <u>Links: Mittlere Zeit in Prozent, die belgische <i>Trainees</i> auf <i>E-learning</i>, Klassenunterricht und Simulatortraining verwendeten Rechts: Mittlerer Lerneffekt (Nutzen) in Prozent, den die belgischen <i>Trainees</i> <i>E-learning</i>, Klassenunterricht und Simulatortraining zuwiesen</u> | 120 |
| <u>Abbildung 57:</u> | <u>Mittlere Anzahl Klassen- (oben) und Simulatorstunden (unten), die <i>E-learning</i> in der schwedischen Stichprobe ersetzt hat</u> | 121 |
| <u>Abbildung 58:</u> | <u>Wahrgenommener, mittlerer Nutzen der vier schwedischen <i>E-learning</i> Module</u> | 126 |
| <u>Abbildung 59:</u> | <u>Wahrgenommener, mittlerer Nutzen der vier belgischen <i>E-learning</i> Module</u> | 127 |
| <u>Abbildung 60:</u> | <u>Prozent erfüllte Qualitätskriterien der schwedischen und belgischen <i>E-learning</i> Module</u> | 127 |
| <u>Abbildung 61:</u> | <u>Persönlicher Nutzen und empfohlener Ausbildungsanteil von <i>E-learning</i> in Schweden und Belgien</u> | 128 |
| <u>Abbildung 62:</u> | <u>Strukturmodell zur ökonomischen Evaluation</u> | 135 |
| <u>Abbildung 63:</u> | <u>Das erweiterte HPM als differenziertes Prozess- und Strukturmodell</u> | 139 |

5.2 Tabellen

| | | |
|-------------------|--|-----|
| <u>Tabelle 1:</u> | <u>Gliederung der Handlungshilfe</u> | 17 |
| <u>Tabelle 2:</u> | <u>Mittlere Anzahl M von Simulator- und Klassenstunden, die durch die besten <i>E-learning</i> Module ersetzt wurden (Angaben von 8 schwedischen <i>Trainees</i> bei 29 <i>Trainees</i> insgesamt).</u> | 124 |
| <u>Tabelle 3:</u> | <u>Mittlere Anzahl M von Simulator- und Klassenstunden, die durch die besten <i>E-learning</i> Module ersetzt wurden (Angaben von 8 belgischen <i>Trainees</i> bei angenommenen 24 <i>Trainees</i> insgesamt).</u> | 125 |

6 Formelzeichen und Abkürzungen

| | |
|--------------|--|
| AIS | Allgemeines Informationssystem |
| ANSP | Air Navigation Service Provider |
| BAuA | Bundesanstalt für Arbeitsschutz und Arbeitsmedizin |
| BEP | Break Even Point |
| CISM | Critical Incident Stress Management |
| COS | Cief of Section, früher Seniorwachleiter in der Flugsicherung |
| DFS | Deutsche Flugsicherung |
| EBG | Einsatzberechtigungsgruppe, früher Wache in der Flugsicherung; operative Einheit bei der DFS |
| EMS | Economic Method Selection Tool |
| FAQ | Frequently Asked Questions |
| HERA | Human Error Analysis |
| HF | Human Factors |
| HPM | <i>Human Resources Performance Model</i> |
| HR | Human Resource |
| HRE | Human Resource Evaluation |
| HRM | Human Resource Management |
| IANS | Institute of Air Navigation Services |
| ICISF | International Critical Incident Stress Foundation |
| KVP | Kontinuierlicher Verbesserungsprozess |
| NL | Niederlassung |
| Nutzwert | Siehe UA |
| <i>Peers</i> | Fluglotsen, die einen Kollegen nach einem kritischen Vorfall betreuen |
| ROI | Return on Investment |
| Rubikon | Fluss, der etwa 40 Kilometer nordöstlich von Florenz in den mittleren Höhenlagen der Apenninen entspringt. Bekannt wurde der Ru- |

| | |
|----|--|
| | <p>bikon ab 49 v. Chr. durch den römischen Bürgerkrieg (Gaius Julius Caesar gegen Gnaeus Pompeius Magnus). Als der Römische Senat beschloss, dass Caesar sein Heer entlassen müsse, überschritt dieser den Grenzfluss Rubikon mit der Bemerkung, dass die Würfel geworfen seien. Er eroberte daraufhin ganz Italien und später Spanien. Heute steht der Ausdruck „den Rubikon überschreiten“ für das unwiderrufliche Involvierem in eine Handlung. Der Ausdruck stand Pate für das psychologische Handlungsmodell von Heinz Heckhausen, das den Übergang vom Wollen zum Tun beschreibt. Im vorliegenden Bericht wird der Ausdruck psycho-ökonomischer Rubikon als Synonym für den entscheidenden Schritt benutzt, Psychologie und Ökonomie miteinander zu verbinden und der strategischen und ökonomischen Bedeutung des so genannten <i>Soft Factors</i> Personal Rechnung zu tragen.</p> |
| SE | Standard error of the mean |
| SV | Supervisor, früher Wachleiter in der Flugsicherung |
| T | Training |
| UA | Utility Analysis |

Anh. 1: CISM Interviewleitfaden Manager¹⁷

Anh. 2: CISM Lotsenfragebogen

Anh. 3: CISM Peerfragebogen

Anh. 4: PE-Programm Mitarbeiterfragebogen

Anh. 5: PE-Programm Führungskräftefragebogen

Anh. 6: *E-learning* Interviewleitfaden Manager

Anh. 7: *E-learning* Traineefragebogen

Anh. 8: *E-learning* Trainerfragebogen

¹⁷ Alle Anhänge sind auf Anfrage hin bei den Autoren verfügbar.