

V&R **unipress**

Applied Research in Psychology and Evaluation

Vol. 3

Edited by / Herausgegeben vom:
Zentrum für Evaluation und Methoden (ZEM)
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Prof. Dr. Georg Rudinger, Direktor



Georg Rudinger / Katharina Hörsch /
Thomas Krüger (Hg.)

Forschung und Beratung –
Das Zentrum für Evaluation und
Methoden

V&R unipress
Bonn University Press



„Dieses Hardcover wurde auf FSC-zertifiziertem Papier gedruckt. FSC (Forest Stewardship Council) ist eine nichtstaatliche, gemeinnützige Organisation, die sich für eine ökologische und sozialverantwortliche Nutzung der Wälder unserer Erde einsetzt.“

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-89971-761-7

Veröffentlichungen der Bonn University Press erscheinen im Verlag V&R unipress GmbH.

© 2009, V&R unipress in Göttingen / www.vr-unipress.de

Alle Rechte vorbehalten. Das Werk und seine Teile sind urheberrechtlich geschützt. Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf der vorherigen schriftlichen Einwilligung des Verlages. Hinweis zu § 52a UrhG: Weder das Werk noch seine Teile dürfen ohne vorherige schriftliche Einwilligung des Verlages öffentlich zugänglich gemacht werden. Dies gilt auch bei einer entsprechenden Nutzung für Lehr- und Unterrichtszwecke. Printed in Germany.

Titelbild: PhotoCase.com

Gedruckt auf alterungsbeständigem Papier.

Inhalt

VORWORT	9
GEORG RUDINGER Das Zentrum für Evaluation und Methoden	11
TEIL I EVALUATION UND QUALITÄTSSICHERUNG IM UNIVERSITÄREN KONTEXT	
GEORG RUDINGER Hochschulautonomie als Change Prozess: Der Weg zum umfassenden Qualitätsmanagement	25
KATHARINA HÖRSCH & GEORG RUDINGER Qualitätssicherung von Lehre und Studium an Hochschulen: Von Evaluationsergebnissen zu Maßnahmen – aktuelle Entwicklungen an der Universität Bonn	39
RENÉ COSTA Das Bonner Modell der Hochschulevaluation: Erste Erfahrungen mit der Implementierung einer webbasierten Modul- und Lehrveranstal- tungsevaluation.....	49
EVA SINDERN & SANDRA PIETRANGELI Validitätsforschung im Zuge der Entwicklung fachspezifischer Online Self-Assessments an der Rheinischen Friedrich-Wilhelms- Universität Bonn	55
MAXIMILIAN PONERT & KATHARINA HÖRSCH Evaluation von Change Management am Beispiel der Bonner Universitätsverwaltung.....	67
KATHARINA HÖRSCH & GEORG RUDINGER Der Net Promoter Score® an der Universität Bonn	75

ALEXANDRA BORSCH
Der Net Promoter Score (NPS) als diagnostisches Instrument zur
Analyse der Loyalität interner Kunden..... 89

TEIL II EVALUATION UND QUALITÄTSSICHERUNG IN VERSCHIEDENEN
ANWENDUNGSFELDERN

KRISTINA HEIDEMANN
Perspektiven zur Evaluation von Entsendungsprozessen und
Auslandseinsätzen – die Entwicklung eines längsschnittlichen
Designs 107

MAXIMILIAN PONERT & UWE KLEINEMAS
Wissenschaftliche Museumsevaluation: Ein Bericht erfolgreicher
Kooperationen zwischen dem Zentrum für Evaluation und
Methoden und Museen der Bonner Region 127

GEORG RUDINGER, KIERAN DONAGHY & STEFAN POPPELREUTER
Societal trends, mobility behaviour and sustainable transport in
Europe and North America..... 133

GEORG RUDINGER & KRISTINA KOCHERSCHIED
Verkehrssicherheitsperspektiven für eine in die Jahre kommende
Gesellschaft 153

KRISTINA KOCHERSCHIED
Konzeption und Evaluation einer ärztlichen Fortbildung zur
Mobilitätsberatung älterer Kraftfahrer..... 181

GEORG RUDINGER & ANTON REITER
Zur Evaluation der theoretischen Fahrerlaubnisprüfung 193

RENATE SOELLNER, STEFAN HUBER, NORBERT LENARTZ & GEORG
RUDINGER
Gesundheitskompetenz..... 205

TEIL III UMFragen

THOMAS KRÜGER & GEORG RUDINGER
Rektoren-Ranking: Rektor des Jahres 2009..... 223

THOMAS KRÜGER & ASTRID MAYERBÖCK Der Net Promoter Score als Kennziffer zur Unternehmenssteuerung	231
MICHAEL PAULUS Using NPS to play offense	241
ASTRID MAYERBÖCK Bedarfsanalyse eines Gründerzentrums an der Universität Bonn	247
UWE FEESER-LICHTERFELD, SANDRA PIETRANGELI, EVA SINDERN, UWE KLEINEMAS, MICHAEL FUCHS & GEORG RUDINGER Lebensverlängerung und Verlangsamung des menschlichen Alterns	257
 TEIL IV METHODEN UND STATISTIK	
NORBERT HILGER & THOMAS KRÜGER Zur Güte von Hörererhebungen in der Radioforschung	273
GEORG RUDINGER Analyse von Daten aus Längsschnittstudien: Eine Einführung	285
SIMON RIEK, THOMAS KRÜGER & GEORG RUDINGER Befragungsmethodische Aspekte des IW-Zukunftspanels	297
 TEIL V PERSPEKTIVE	
HEIKE STECKHAN, KATRIN WAHLERS & KRISTINA HEIDEMANN Wege zur Professionalisierung: Masterstudiengang Evaluation	319
UMFRAGEPROJEKTE DES ZEM VON A BIS Z	325
AUTORENVERZEICHNIS	337

Vorwort

Das Zentrum für Evaluation und Methoden (ZEM) als universitäre Einrichtung arbeitet in einem energiereichen Kraftfeld konstituiert aus Wissenschaft – Forschung und Lehre – einerseits, Dienstleistung – Auftragsforschung und Beratung – andererseits. Seit der Gründung des ZEM im Jahre 1999 wurden verschiedenste Projekte durchgeführt, bei denen die im ZEM gebündelten Kompetenzen in Forschung, Beratung und Durchführung von Umfragen in vielfältiger Weise zur Geltung gekommen sind. Als inhaltliche Schwerpunkte haben sich über die Zeit die Arbeitsbereiche Evaluation und Qualitätssicherung, Umfragen und Statistik sowie Methodenentwicklung und -anwendung herausgebildet. In diesen Bereichen verfügt das ZEM über langjährige Erfahrung, erworben und vertieft durch Kooperationen mit zahlreichen Partnern und Auftraggebern aus Wissenschaft und Wirtschaft. Entsprechend ist auch dieser Band gegliedert: Er beginnt mit Darstellungen von Evaluations- und Qualitätssicherungsprojekten im universitären Bereich. Es folgt die Schilderung von Projekten in verschiedenen Anwendungsfeldern. In einem dritten und vierten Teil werden Ergebnisse von Umfragen zu verschiedenen Sujets und Methodenentwicklungen wiedergegeben. Zum Abschluss wird kurz der berufsbegleitende Masterstudiengang Evaluation vorgestellt, der natürlich im Kontext der vorgenannten Aktivitäten zu sehen ist.

An dieser Stelle ist auch Dank auszusprechen, Dank all den Freunden und Partnern beginnend beim Innovationsministerium des Landes Nordrhein-Westfalen über die Universität Bonn – Rektorat, Verwaltung, Fakultäten, manche Institute und Zentren sind hier zu nennen – und nicht zuletzt Dank an alle außeruniversitären Kooperationspartner, von denen sich viele in den Projektdarstellungen wieder finden.

Anlässlich des zehnjährigen Bestehens des ZEM gibt der vorliegende Band der Reihe *Applied Research in Psychology and Evaluation* also einen Einblick in die breit gefächerten Projekte, welche – so hoffen wir – eine solide Basis für zukünftige Unternehmungen bilden und günstige Perspektiven für die nächsten zehn Jahre eröffnen.

Bonn, im September 2009

Die Herausgeber

Georg Rudinger

Das Zentrum für Evaluation und Methoden¹

Das Zentrum für Evaluation und Methoden (ZEM) wurde 1999 von Prof. Dr. Georg Rudinger an der Philosophischen Fakultät gegründet. Neben Forschungsaktivitäten in den Bereichen Evaluation und Qualitätssicherung, Umfragen und Statistik besteht die Hauptaufgabe des ZEM in der Durchführung der Evaluations- und Qualitätssicherungsmaßnahmen an der Rheinischen Friedrich-Wilhelms-Universität Bonn (Rektoratsbeschluss vom 13. 4. 2005). Das ZEM bietet somit Service für Universität und Fakultäten im Bereich der Qualitätssicherung und Evaluation, betreibt grundlagen- und anwendungsorientierte Forschung mit Partnern aus dem außeruniversitären Umfeld und bietet seit dem Sommersemester 2008 einen gebührenpflichtigen Weiterbildungsstudiengang Evaluation und Qualitätssicherung an.

1. Forschung – Grundlagen, Anwendung und Beratung

Die aktuellen Aktivitäten des ZEM in Forschung und Beratung lassen sich drei Bereichen zuordnen

– Evaluation und Qualitätssicherung

In diesem Bereich können die Forschungs-/Aktivitäten des ZEM zur *Evaluation kultureller und kulturpolitischer Programme* ebenso positioniert werden wie die zur *Evaluation und Qualitätssicherung von Bildungsprogrammen* in der Aus-, Weiter- und Fortbildung. Nicht zuletzt ist hier die Funktion und Stellung des ZEM als »Rektoratsstabstelle« und Dienstleistungseinheit für die Universität Bonn zu nennen (s. u.). Im Schwerpunkt *Verkehr, Altern und Gesundheit* befasst sich das ZEM mit Programmen der schulischen und weiterbildenden Verkehrserziehung sowie der Evaluation von Programmen der Verkehrsaufklärung vor allem von Senioren (Kooperation mit der Bundesanstalt für Straßenwesen – BASt).

– Umfragen und Statistik

Das ZEM führt im universitären und außeruniversitären Bereich im Rahmen von Drittmittelprojekten umfragegestützte Forschung zu verschiedenen Themenfeldern durch. Dies geschieht mit und für Partner/n wie Institut der deutschen Wirtschaft (IW Köln), Zentrum für europäische Wirtschaftsforschung (ZEW), Creditreform, Impuls/TÜV Rheinland, Stifterverband für die

¹ Auszug aus dem Struktur- und Entwicklungsplan des ZEM (2009)

deutsche Wissenschaft, Accenture, IBM Business Consult, Wissenschaftsrat (WR), Deutscher Hochschulverband (DHV).

– Entwicklung und Anwendung statistischer Methoden

Das dritte Forschungsfeld, die statistischen Methodenforschung, geht zunächst von dem Problem aus, dass (vor allem bei Umfragen) häufig selektive Stichproben mit zahlreichen fehlenden Werten untersucht werden, deren Auswertung mit gängigen statistischen Verfahren wegen Voraussetzungsverletzungen oder geringer Fallzahlen nicht oder nur eingeschränkt möglich ist (u. a. intensive Zusammenarbeit mit dem IW Köln: Zukunftspanel) und schließt Panel- und Längsschnittanalysen mit multivariaten Verfahren der zweiten und dritten Generation ein (Strukturgleichungsmodelle/Structural Equation Modeling – SEM, Hierarchical Linear Modeling – HLM).

1.1 Evaluation und Qualitätssicherung

Der Schwerpunkt *Kultur und Kulturpolitik* hat sich aus der Arbeitsgruppe »Expatriates – Auslandsentsendung mit System« des ZEM entwickelt. Ziel dieser Arbeitsgruppe war es, Forschung und Praxis im Kontext der Auslandsentsendung zum gegenseitigen Nutzen sinnvoll miteinander zu verbinden. Die Arbeitsgruppe setzte sich mit den Schwierigkeiten der Auslandsentsendung von Mitarbeitern, ihren Angehörigen und Kollegen in Profit- und Non-Profit-Organisationen auseinander. Dabei standen ein ganzheitliches Konzept zur Bewertung von Auslandseinsätzen und den begleitenden Maßnahmen sowie die Berücksichtigung aller beteiligten Akteure im Vordergrund. Zur Evaluation entsprechender Maßnahmen wurden am ZEM längsschnittliche Szenarien entwickelt und erprobt. Die Forschungsergebnisse der Arbeitsgruppe wurden in einer Schriftenreihe veröffentlicht.

Im Feld *Kultur und Kulturpolitik* führt das ZEM heute u. a. Evaluationen von Aus- und Weiterbildungsprogrammen vor allem aus der auswärtigen Kulturpolitik durch. Ein Schwerpunkt liegt dabei auf der Evaluation von Förderprogrammen von Stiftungen wie der *Alexander von Humboldt-Stiftung* oder der *Robert Bosch Stiftung* und anderen Trägern (wie z. B. des *Auswärtigen Amtes*). Des Weiteren fallen darunter Verbleibstudien von internationalen Programmen (z. B. *ILT Tracer Studies für Inwent* – Internationale Weiterbildung und Entwicklung gGmbH).

Zu erwähnen sind hier auch die langjährigen Kooperation mit der *Bundeskunsthalle* (KAH) und dem *Rheinischen Landesmuseum* (LVR).

Im Bereich *Verkehr, Altern und Gesundheit* befasst sich das ZEM mit Programmen der schulischen und weiterbildenden Verkehrserziehung sowie der Evaluation von Programmen der Verkehrsaufklärung vor allem von Senioren. Beispielhaft genannt seien hier die folgenden Projekte: Entwicklung von Evaluationsinstrumenten zur Erreichung von Standards in der Verkehrs-/Mobilitätserziehung der Sekundarstufe (*Bundesanstalt für Straßenwesen* –

BASSt), Ältere Menschen im Straßenverkehr (BASSt), Freizeitmobilität älterer Menschen (BMBF), Verkehrssicherheitsbezogene Botschaften für Senioren (BASSt), Entwicklung und Evaluation eines Mobilitäts-Screenings zur Erfassung der Fahrkompetenz älterer Kraftfahrer (BASSt), Sustainable Transport in Europe and Links and Liaisons with America (*Europäische Union*), Subjektive Sicherheitswahrnehmung im Verkehr (BASSt).

Im Bereich *Bildung* mit ihren Facetten *Aus-, Weiter- und Fortbildung* ist das ZEM verantwortlich für die Umsetzung der Hochschulevaluation der Universität Bonn. Die regelmäßige *Evaluation* dient der systematischen Analyse, Sicherung und Verbesserung der Qualität von Lehre und Studium. Die aus der Evaluation hervorgehenden Informationen dienen zugleich der internen und externen Rechenschaftslegung und stellen eine wesentliche Grundlage für die (Re-) Akkreditierung von Studiengängen dar. Wesentliches Element der internen Evaluation von Studium und Lehre nach dem Bonner Modell der Hochschulevaluation ist die onlinegestützte Datenerhebung unter Einbeziehung aller beteiligten Akteure (i. W. Studenten, Lehrende und Absolventen), welche somit an der Ermittlung des Ist-Zustands, der Ableitung von Lehrzielen und Maßnahmen sowie deren Umsetzung mitwirken. Die Evaluation nach dem Bonner Modell umfasst neben einer allgemeinen Studierendenbefragung die regelmäßige Lehrveranstaltungs- und Modulevaluation sowie die Absolventenbefragung. Das Bonner Modell zur Hochschulevaluation wurde an der Rheinischen Friedrich-Wilhelms-Universität als Regelverfahren, verankert in der Evaluationsordnung, eingeführt und steht auch anderen Universitäten zur Verfügung. Beispielsweise kommt das Bonner Modell an der Universität Leipzig und an der Alanus Hochschule Alfter zum Einsatz.

Evaluation und Qualitätssicherung von Studium und Lehre stellt ein »Modul« im Gesamtprozess der *Bindung* von Studierenden an ihre Universität dar. Dieser Prozess beginnt mit der Gewinnung von Schülern und erstreckt sich bis in die Alumni-Arbeit. Neben zahlreichen Ansätzen für die erste Phase ist hier die Konzeption und Konstruktion eines *Online-Self-Assessment Verfahrens (OSA)* zu nennen, welches vom ZEM in Zusammenarbeit mit einer eigens vom Rektorat eingerichteten Lenkungsgruppe und unter engagierter Beteiligung der ausgewählten Fächer und Studiengänge erstellt wird und für einige Fächer bereits angeboten wird.

Das ZEM fungiert darüber hinaus als *Dienstleistungseinheit der Universität Bonn*. Es übernimmt interne Evaluationen von Maßnahmen universitärer Einrichtungen oder unterstützt sie bei der Selbstevaluation von Maßnahmen. Die vom ZEM durchgeführten internen Evaluationen finden meist programmbegleitend statt und dienen der Maßnahmensteuerung. Sie erfassen den aktuellen Erfolg der laufenden Projekte bzw. Programme und stellen Optimierungspotentiale heraus. Abhängig vom Umfang der beauftragten Evaluation nimmt das ZEM mit Abschluss der Maßnahmen auch eine Beur-

teilung des Erfolgs der Programme bzw. Projekte vor. Entsprechend der Zielstellungen der Evaluation und der definierten Indikatoren setzt das ZEM unterschiedliche Methoden zur quantitativen und qualitativen Datenerhebung ein. Es handelt sich dabei unter anderem um folgende Evaluationen:

- Evaluation des Einführungsprozesses von M.O.S./e-M.O.S. in der Verwaltung der Universität Bonn
- Evaluation des »Sehen, Hören, Verstehen« Medienkonzepts der Sonderforschungsbereiche der Universität Bonn
- Evaluation des Mentoring- und Trainingsprogramm des Gleichstellungsbüros der Universität Bonn
- Evaluation des Betreuungsangebots für ausländische Studierende der Universität Bonn sowie die Evaluation verschiedener weiterer Programme und Projekte der Universität Bonn.

In diesem Zusammenhang ist das Projekt *Hochschulautonomie als Change Prozess: Der Weg zum umfassenden Qualitätsmanagement* zu nennen, welches das ZEM im Auftrag und in enger Zusammenarbeit mit dem Rektorat dem Stifterverband im Jahre 2007 (Präsentation als eines von 12 aus 68 Projekten im Januar 2008 in Berlin) vorgelegt hatte. Der Stifterverband sah den Anlass für eine solche Ausschreibung darin, dass durch staatliche Steuerung und Akkreditierung die Qualitätssicherung an Hochschulen lange Zeit von außen bestimmt worden ist. Autonome Hochschulen sind für die Qualität ihrer Leistungen jedoch selbst verantwortlich. Aus dieser Verantwortung heraus etablieren sie funktionierende Systeme zur Qualitätssicherung und -verbesserung aller Leistungsprozesse. Wenn Hochschulen solche Systeme eingeführt haben, können sich Staat und Akkreditierungsagenturen weitgehend auf die Überprüfung der Funktionsfähigkeit des Qualitätsmanagements beschränken.

Ziel des beantragten Projektes war (und ist) die Entwicklung und Implementierung eines universitätsweiten Steuerungssystem zum Qualitätsmanagement (QM) auf der Basis von qualitätsrelevanten Kennziffern als strukturierendes und zugleich treibendes Element für den Veränderungsprozess. Dieses Steuerungssystem integriert die Bereiche Forschung, Lehre und Management und trägt zu einer weiteren Profilierung und sichtbaren Positionierung der Universität Bonn bei. Die Einführung dieses übergeordneten steuerungsrelevanten Systems wird als Change Management Prozess im Sinne der professionellen Gestaltung von Veränderungsprozessen in Organisationen verstanden.

1.2 Umfragen und Statistik

Im Rahmen der so genannten »kleinen Zielvereinbarung« mit dem MIWFT NRW wurde ein Umfragezentrum eingerichtet, mit dessen Hilfe umfangreiche Fragebogen- oder Telefonuntersuchungen sowie onlinegestütz-

te Erhebungen durchgeführt werden können. Häufig gibt es im sozial- und wirtschaftswissenschaftlichen Kontext Befragungen, die ein großes Maß an Sensibilität bzw. Vertrauen schaffende Maßnahmen erfordern (z. B. Fragen zu persönlichen Finanzverhältnissen oder geschäftlichen Strategien). Um auch schwierige Zielgruppen kontaktieren zu können, entwickelt das ZEM in Zusammenarbeit mit externen Auftraggebern neue Strategien und Methoden inhaltlich vertrauliche Befragungen mit verschiedenen Befragungsmedien durchzuführen. Zusätzlich sollen diese auch an Befragungen im Sinne von Panel- bzw. Längsschnittstudien gebunden werden. Beispiele für solche Zielgruppen sind niedergelassene oder in Kliniken arbeitende Ärzte, Geschäftsführer, Personalverantwortliche im Profit- und Non-Profit-Bereich, Leiter von Forschungs- und Entwicklungsabteilungen. Durch die Kombination verschiedener Ansätze und Techniken konnte ein Methodenmix entwickelt werden, der als Ansatz zur Lösung des Problems der schwierigen Erreichbarkeit verstanden werden kann. Dieser Methodenmix hat sich inzwischen in der Praxis bewährt und wird kontinuierlich weiterentwickelt.

In der Umfrageforschung wird die gleichzeitige Anwendung verschiedener Erhebungsmethoden nach wie vor kontrovers diskutiert. Das ZEM untersucht daher mittels theoretischer und empirischer Studien, welche Auswirkungen verschiedene Befragungsmethoden auf die Validität empirischer Ergebnisse haben und unter welchen Bedingungen methodenabhängige Verzerrungen zu erwarten sind. Speziell die Methode der Online-Umfrage ist häufig umstritten, da nicht von einer tatsächlichen Repräsentativität ausgegangen werden kann. Am ZEM werden hierzu Methoden entwickelt, um Online-Umfragen möglichst repräsentativ zu gestalten und Verzerrungsfaktoren zu eliminieren. Dies geschieht erneut durch die Kombination verschiedener Methoden (z. B. telefonische Akquise und anschließende Online-Befragung) sowie durch verschiedene Ansätze des Einsatzes von Remindern.

Die Mitarbeiter des ZEM sind mit längsschnittlichen bzw. Panel-Untersuchungen intensiv vertraut. Zurzeit werden für Spezialpopulationen (z. B. Bonner Studenten, junge Unternehmen) Panel aufgebaut, die in bestimmten Abständen regelmäßig befragt werden. Entwickelt und verfeinert werden Methoden, um die Panelmortalität zu senken und die Repräsentativität auch über längere Zeiträume zu gewährleisten.

Das ZEM versteht sich als auch hier Dienstleister im universitären und außeruniversitären Bereich. Dadurch kann die Bearbeitung von Forschungsfragen mit Projekten verschiedener externer Auftraggeber sinnvoll zu einem umfassenden wissenschaftlich fundierten und gleichzeitig anwendungsbezogenen Tätigkeitsfeld ergänzt werden. Demnach führt das ZEM zusätzlich zu den genannten Forschungsschwerpunkten im Rahmen von Drittmittelprojekten Auftragsforschung zu verschiedenen Themenfeldern durch. Das ZEM erstellt für Auftraggeber aus Wissenschaft und Wirtschaft bereits seit mehreren Jahren regelmäßig Studien und Datensätze. So zählen z.B. das *Institut der*

deutschen Wirtschaft (IW Köln), das Zentrum für europäische Wirtschaftsforschung (ZEW), die Creditreform, TÜV Rheinland bzw. Impuls GmbH, der Stifterverband für die deutsche Wissenschaft, Accenture, IBM Business Consult, der Wissenschaftsrat (WR), das Institut für Forschungsinformation und Qualitätssicherung (IFQ), mit dem ein Kooperationsvertrag besteht, der Deutsche Hochschulverband (DHV), die Stadt Bonn und andere zu unseren Partnern. Insgesamt wurden in den letzten drei Jahren mehr als 200 Erhebungen vom ZEM durchgeführt.

1.3 Methoden: Entwicklung und Anwendung

Vor allem die Aktivitäten des Umfragezentrums sind eng verknüpft mit dem dritten Forschungsfeld, der statistischen Methodenforschung, da häufig selektive Stichproben mit zahlreichen fehlenden Werten untersucht werden, deren Auswertung mit gängigen statistischen Verfahren wegen Voraussetzungsverletzungen oder geringer Fallzahlen nicht oder nur eingeschränkt möglich ist.

Die im Bereich der Methodenforschung geplanten Projekte werden seit 2006 vor allem im Rahmen von Dissertationsarbeiten durchgeführt. Im Bereich der computerintensiven statistischen Verfahren wurde im Herbst 2008 eine umfangreiche Simulationsstudie vorgelegt, die sich u.a. mit verschiedenen Bootstrap-Schätzern von Effektgrößen auseinandersetzt. Resampling-Techniken finden ferner ihren Einsatz in der Begleitforschung von z. B. Projekten zur Markenstärke (*T-Mobile International GmbH*) oder in der Hörverhaltensforschung (*Medialysis GmbH*).

Fragen zur Verzerrung von Umfrageergebnissen aufgrund fehlender Werte und Fälle werden in Zusammenarbeit mit dem *IW Köln* im Rahmen des *IW-Zukunftspanels* behandelt. Eine weitergehende Untersuchung zu statistischen Eigenschaften und Anwendungsaspekten verschiedener Imputationsalgorithmen wird derzeit im Rahmen intensiver Kooperation mit dem *IW Köln* u. a. in weiteren Qualifikationsarbeiten durchgeführt.

Lineare Strukturgleichungsmodelle (SEM) und Mehrebenen-Analyse (HLM) finden im Zusammenhang mit europaweiten bzw. globalen Umfragestudien, die von ZEM-Mitarbeiter/inne/n in Kooperation mit der *T-Mobile International AG* und *IBM Business Consult* geleitet wurden, verstärkte Anwendung.

2. Studium und Lehre

Weiterbildender Masterstudiengang Evaluation

Neben der Anwendung von Evaluation in den oben beschriebenen drei Schwerpunktfeldern ist das ZEM auch als Bildungsanbieter im Bereich Evaluation aktiv. Seit dem Sommersemester 2008 führt das ZEM einen weiter-

bildenden Masterstudiengang Evaluation durch, der im August 2008 durch AQAS akkreditiert wurde. Der berufsbegleitende Studiengang richtet sich an Hochschulabsolventen, die bereits über erste berufliche Erfahrungen im Bereich Evaluation verfügen. Im Verlauf von vier Semestern werden ihnen eine breite theoretische und methodische Wissensgrundlage und darauf aufbauende Praxiserfahrungen vermittelt, die dazu beitragen, Evaluationen zielgerichtet und genau durchzuführen, zu beauftragen und zu begleiten.

Im Wintersemester 2009/2010 wird die erste Kohorte den Studiengang abschließen. Im Sommersemester 2009 nimmt die zweite Kohorte ihre Ausbildung auf. Vor dem Hintergrund der hohen Anzahl an Interessenten ist eine kontinuierliche Weiterführung des Studiengangs geplant. Daneben sollen in Zukunft einzelne Module des Studiengangs auch externen Interessenten gegen entsprechende Gebühr zugänglich gemacht werden, die diese nach erfolgreicher Teilnahme mit einem Zertifikat abschließen.

Der Studiengang wurde anhand der Empfehlungen für die Aus- und Weiterbildung in der Evaluation der Gesellschaft für Evaluation (DeGEval) entwickelt. Durch aktive Mitgliedschaft im Arbeitskreis Aus- und Weiterbildung in der Evaluation der DeGEval werden die kontinuierliche Qualitätssicherung sowie der Austausch mit anderen Anbietern ähnlicher Programme gewährleistet. Es konnten renommierte Dozenten aus der »Evaluation-Community« gewonnen werden. Die Auswahl an Vertiefungsmöglichkeiten wird entsprechend der Entwicklungen des Evaluationssektors und der Interessen der Studierenden angepasst werden.

Statistik und Methoden

Das ZEM exportiert Lehre in den Bereichen Statistik, Methoden und Evaluation in die Philosophische Fakultät (Institut für Psychologie) der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Betreuung von Diplomarbeiten und Dissertationen

Allein in den letzten drei Jahren wurden in Betreuung von Mitarbeitern des ZEM insgesamt 14 Diplomarbeiten fertiggestellt. Zudem wurden in Erstbetreuung von Prof. Rudinger 3 Mitarbeiter des ZEM und 4 Auswärtige mit Arbeiten in den genannten ZEM-Forschungsbereichen promoviert. Aktuell sind 3 Diplomarbeiten und 12 Dissertationen in Bearbeitung.

Theorie und Praxis der Umfrageforschung

Das Zentrum für Evaluation und Methoden hat im Rahmen der Vermittlung von *Schlüsselkompetenzen* in den Bachelor-Studiengängen das Modul »Theorie und Praxis der Umfrageforschung« – eingebettet in das Lehrangebot der Zentralstelle für Schlüsselkompetenzen: ZSK – übernommen.

Im Mittelpunkt des Moduls steht die praktische Durchführung von telefonischen Umfragen. Diese wird durch eine Schulung vorbereitet und durch Feedback-Gespräche sowie durch Supervision begleitet und ausgewertet. Zudem wird die Praxis der Supervision eingeführt. Es sollen die Grundlagen der empirischen Sozial-, Markt- und Meinungsforschung vermittelt werden. Daneben sammeln die Studierenden in dem Modul praktische Erfahrung in diesem Bereich. Zu den Inhalten gehören quantitative und qualitative Umfragen, Umfragearten sowie Qualitätssicherung in der Umfrageforschung.

Betreuung von Feldforschungsprojekten

Das ZEM unterstützt regelmäßig laufende Forschungsprojekte im Rahmen von Lehrveranstaltungen und Forschungsarbeiten für die verschiedenen Fakultäten. So finden regelmäßig Lehrveranstaltungen anderer Institute/Fakultäten unter Rückgriff auf die Datenerhebungsmöglichkeiten im ZEM statt. Darüber hinaus werden diese Institute/Fakultäten z. B. bei eigenen Evaluationsvorhaben unterstützt und in Zusammenarbeit mit dem ZEM umgesetzt.

Export von Methodenkompetenz

Forschungsprojekte verschiedener Institute und Fakultäten werden durch das ZEM regelmäßig betreut und supervidiert (z.B. Institut für Landtechnik/Haushaltstechnik, Institut für Tierhygiene, Evolutionsbiologie, Zahnmedizin, Katholische Hochschulgemeinde). Die Expertise des ZEM wird hier sowohl im Bereich der Planung empirischer Untersuchungen als auch der Auswertung und Ergebnisdarstellung in Anspruch genommen.

Betreuung von Praktikanten in den Forschungsschwerpunkten des ZEM

Durch das ZEM werden Praktikanten an alle Forschungsschwerpunkte herangeführt und fachkundig betreut (durchschnittlich parallel fünf Praktikanten). Studierende erhalten so die Möglichkeit, neben einer fachlich vertiefenden Ausbildung auch Fachqualifikationen im Sinne der sog. Schlüsselkompetenzen zu erwerben.

3. Organisation, Management und Ressourcenausstattung

Das ZEM nimmt – wie beschrieben – im Auftrag des Rektorats der Rheinischen Friedrich-Wilhelms-Universität Bonn z. Zt. regelmäßig folgende zentralen Tätigkeiten wahr:

- Kontinuierliche Erfassung des Absolventenverbleibs,
- Qualitätssicherung und Evaluation in den Bachelor- und Masterstudiengängen und

- Durchführung von Studierendenbefragungen mit zusätzlicher Berücksichtigung von Erstsemestern und Studienortwechslern.
- Entwicklung eines Online Self-Assessment-Instruments (OSA) für ausgewählte Fächer und Studiengänge in Zusammenarbeit mit einer entsprechenden vom Rektorat eingesetzten Lenkungsgruppe mit der Zielsetzung umfassenden OSAs.

Weitere Dienstleistungen für die Universität werden z. B. in den Bereichen

- Mitarbeiterbefragung,
- Mitarbeit an empirischen Erhebungen (z. B. CEWS, KHG, Gleichstellungsbeauftragte) sowie
- -Methodenberatung

erbracht. Die Durchführung dieser Aufgaben ist eng an das ZEM-eigene Umfragezentrum gekoppelt.

Das Umfragezentrum ist darüber hinaus als unabdingbare strukturelle und methodische Voraussetzung für die Forschungstätigkeiten in den beschriebenen Schwerpunkten zu betrachten. Allein in den Jahren 2006 bis 2008 sind mit und für diese/n Forschungen ca. 2,3 Mio. Euro eingeworben worden.

Zu erwähnen ist hier auch die Buch-Reihe, welche nach »Akkreditierung« durch den Beirat bei Bonn University Press mit jährlich 1 bis 2 Bänden erscheint: *Applied Research in Psychology and Evaluation* (ARPE).

Arbeitsgruppe und Beirat

Das ZEM wird zur Unterstützung seiner Zentralaufgaben in Evaluation und Qualitätssicherung von einer Arbeitsgruppe aus den Fakultäten begleitet werden.

Darüber hinaus plant das ZEM die Einrichtung eines Beirates mit ausgewiesenen internationalen Experten aus den drei eingangs beschriebenen Forschungsbereichen Evaluation und Qualitätssicherung, Umfragen und Statistik, Entwicklung und Anwendung statistischer Methoden.

4. Ausgewählte Publikationen

ENGIN, T., Kocherscheid, K., Feldmann, M. & Rudinger, G. (2009) *Entwicklung und Evaluation eines Screening-Tests zur Erfassung der Fahrkompetenz älterer Kraftfahrer*. (Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit). Bremerhaven: Wirtschaftsverlag NW.

HEIDEMANN, K. (2009). *Perspektiven zur Evaluation von Auslandseinsätzen – ein längsschnittliches Design*. Saarbrücken: VDM. Zugriff am 17.09.2009 unter http://hss.ulb.uni-bonn.de/diss_online/phil_fak/2008/heidemann_kristina.

HEIDEMANN, K., Steckhan, H. & Rietz, C. (Hrsg.) (2007). *Erfolg von Auslandseinsätzen: Neue Beurteilungsansätze. Einfluss auf die Persönlichkeit und Kriterien zur Erfolgsbeurteilung*. (Schriftenreihe der Arbeitsgruppe »Expatriates« am

- Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn), Bonn: PACE.
- HÖRSCH, K. & Rudinger, G. (2009). Von Evaluationsergebnissen zu Maßnahmen – aktuelle Entwicklungen an der Universität Bonn. In: G. Rudinger. & K. Hörsch (Hrsg.), *Umsetzung von Evaluationsergebnissen: Theorie und Praxis*. (Applied Research in Psychology and Evaluation 2, S. 129-134). Bonn University Press; Göttingen: V&R unipress.
- ILLES, F., Bernhardt, T., Prell, K., Rietz, C., Rudinger, G., Froelich, L., Maier, W., Rietschel, M. (2006). Einstellung zu genetischen Untersuchungen auf Alzheimer-Demenz. *Zeitschrift für Gerontologie und Geriatrie*, 39 (3), 233-239.
- KAPPELHOFF, V., Heidemann, K., Völker, S. & Rietz, C. (2006). Partnerschaft und Auslandseinsatz: Beziehungsdimensionen und ihre Bedeutung für die Qualität des Auslandseinsatzes bei Expatriates. *Zeitschrift für Personalforschung*, 20 (4), 318-342.
- KOCHERSCHIED, K. (2009). *Konzeption und Evaluation einer ärztlichen Fortbildung zur Mobilitätsberatung älterer Kraftfahrer*. Saarbrücken: Südwestdeutscher-Verlag für Hochschulschriften.
- KOCHERSCHIED, K., Rietz, C., Poppelreuter, S., Riest, N., Engin, T., Müller, A. & Rudinger, G. (2007). *Verkehrssicherheitsbotschaften für Senioren: Nutzung der Kommunikationspotenziale im allgemeinmedizinischen Behandlungstag*. (Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit, Heft M 184). Bremerhaven: Wirtschaftsverlag NW.
- KOTTE, S., Heidemann, K., Steckhan, H. & Rietz, C. (2006). *Reintegration nach Auslandseinsätzen: Eine empirische Untersuchung zur effektiven Kommunikation im Ausland gewonnener Erfahrungen*. (Schriftenreihe der Arbeitsgruppe »Expatriates« am Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn). Bonn: PACE.
- KRÜGER, T. & Rudinger, G. (2009). Rektoren-Ranking 2009: Ergebnisse einer Umfrage. *Forschung & Lehre*, 4/09, 260-263.
- RIEK, S., Krüger, T. & Rudinger, G. (2009). Befragungsmethodische Aspekte. In K. Lichtblau & A. Neligan (Hrsg.) *Das IW-Zukunftspanel* (S.35-50). Köln: Deutscher Instituts-Verlag
- RUDINGER, G. & Hörsch, K. (2009, i. Druck). Wirkungen von Qualitätsentwicklung auf das Studium – hat sich die Qualität der Lehre durch Qualitätssicherung wirklich verbessert? In HRK (Hrsg.), *Methoden und Wirkungen von Qualitätssicherung im Europäischen Hochschulraum*. HRK: Bonn.
- RUDINGER, G. & Hörsch, K. (Hrsg.). (2009). Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung. (Applied Research in Psychology and Evaluation 4). Bonn University Press; Göttingen: V&R unipress.
- RUDINGER, G. & Hörsch, K. (Hrsg.) (2009). *Umsetzung von Evaluationsergebnissen: Theorie und Praxis*. Applied Research in Psychology and Evaluation 2. Bonn University Press. Göttingen: V&R unipress.
- RUDINGER, G. & Kocherscheid, K. (2007). *Verkehrssicherheitsperspektiven für eine in die Jahre kommende Gesellschaft*. ADAC/BASi-Symposium am 13. Oktober

- 2006 in Baden-Baden. (Berichte der Bundesanstalt für Straßenwesen: Mensch und Sicherheit, Heft M 191, S.73-95). Bergisch Gladbach: BAST.
- RUDINGER, G. (2006). Technik für alte Menschen am Beispiel der Automobilität. In M. Kerner & T. Müller (Hrsg.), *Gespaltene Welt? Technikzugänge in der Wissensgesellschaft* (S. 213-23). Köln: Böhlau.
- RUDINGER, G. (2009, im Druck). Längsschnittstudien. In H.Holling & B. Schmitz (Hrsg.), *Handbuch der psychologischen Methodenlehre und Evaluation*. Göttingen: Hogrefe.
- RUDINGER, G., Donaghy, K. & Poppelreuter, S. (2006). Societal trends, mobility behaviour and sustainable transport in Europe and North America. *European Journal of Transport and Infrastructure Research*, 6 (1), 61-76.
- RUDINGER, G., Krahn, B. & Rietz, C. (2007). Möglichkeiten und Perspektiven online-gestützter Verfahren im Kontext von nachhaltiger Qualitätssicherung und Evaluation. In R. Keil, M. Kerres & R. Schulmeister (Hrsg.) *e-University – Update Bologna* (S.223-242). Münster: Waxmann.
- RUDINGER, G., Krahn, B., & Rietz, C. (Hrsg.) (2008). *Evaluation und Qualitätssicherung von Forschung und Lehre im Bologna-Prozess*. (Applied Research in Psychology and Evaluation 1). Bonn University Press; Göttingen: V&R unipress.
- SOELLNER, R., Lenartz, N., Huber, S. & Rudinger, G. (2008). Health literacy – a concept map. *Psychology & Health*, 23, (Supplement 1), 240.
- STECKHAN, H., Wahlers, K., Heidemann, K. (2008). Wege zur Professionalisierung: Masterstudiengang Evaluation. *Zeitschrift für Evaluation*, 7 (2), 303-307.

Teil I
Evaluation und Qualitätssicherung im
universitären Kontext

Dieser Teil umfasst sieben Artikel, deren Reihe mit der Skizze des Projekts *Hochschulautonomie als Change Prozess: Der Weg zum umfassenden Qualitätsmanagement*¹ eröffnet wird, gefolgt von der Beschreibung der Hochschulevaluation an der Universität Bonn. *Evaluation und Qualitätssicherung von Studium und Lehre* stellt ein »Modul« im *Bindungsmanagement* dar. Dieser Prozess beginnt mit der Gewinnung von Schülern und erstreckt sich bis in die Alumni-Arbeit. Wesentliches Element ist hier die Konzeption und Konstruktion von *Online-Self-Assessment Verfahren*. Qualität der Bindung lässt sich u. a. durch die Intensität der Weiterempfehlungsbereitschaft über den *Net Promotor Score*® erfassen. Als Dienstleistungseinheit der Universität Bonn übernimmt das ZEM natürlich auch *interne Evaluationen von Maßnahmen* universitärer Einrichtungen – ein Beispiel wird gegeben.

1 *Kursive Passagen* benennen bzw. umschreiben die Themen der Artikel

Georg Rudinger

Hochschulautonomie als Change Prozess: Der Weg zum umfassenden Qualitätsmanagement¹

Umfassendes Qualitätsmanagement an Universitäten integriert die Bereiche Forschung, Lehre, Management und Verwaltung. Der hier beschriebene Ansatz soll zu einer weiteren Profilierung und sichtbaren Positionierung der Universität Bonn beitragen. Die Einführung eines solchen übergeordneten und steuerungsrelevanten Systems wird als Change Management Prozess im Sinne der professionellen Gestaltung von Veränderungsprozessen in Organisationen verstanden. Die Universität Bonn wird ein umfassendes Qualitätsmanagement im Kontext einer auf mittel- bis langfristig wirksame Veränderungen von Verhaltensmustern und Fähigkeiten zielenden Strategie unter ganzheitlicher Betrachtung der »Organisation Universität« implementieren. Gleichzeitig ist die Universität, d. h. deren Führung, der wichtigste Akteur mit institutioneller Eigenverantwortung für die Verwirklichung von Qualitätssicherung und Qualitätsentwicklung, Profilbildung sowie ganzheitlicher Steuerung.

1. Ziele und Konzeption

Den Hochschulen mehr Freiheit, mehr Verantwortung, mehr Autonomie zu gewähren, stellt geradezu einen Paradigmenwechsel dar, der im System Hochschule tief greifende funktionale und strukturelle Veränderungen mit sich bringt bzw. von den Hochschulen als Hauptakteuren selbst die Initiierung solcher Veränderungen verlangt. Sie betreffen die strategische, technologische, strukturelle und personelle Dimension einer jeden Universität und stellen entsprechend einen umfassenden Change Prozess dar.

In einem Zeitraum von zwei Jahren soll als strukturierendes und zugleich treibendes Element für diesen Veränderungsprozess ein universitätsweites Steuerungssystem zum Qualitätsmanagement auf der Basis von qualitätsrelevanten Kennziffern (KPI = Key Performance Indicator) implementiert sein, welches die Bereiche Forschung, Lehre und Management integriert und zu

1 Antrag der Rheinischen Friedrich-Wilhelms-Universität Bonn an die Heinz Nixdorf Stiftung und den Stifterverband für die deutsche Wissenschaft vom 15. Oktober 2007, vorgestellt in Berlin am 11. Januar 2008; Ansprechpartner innerhalb der Universität: Prof. Dr. Georg Rudinger (ZEM)

einer weiteren Profilierung und sichtbaren Positionierung der Universität Bonn beiträgt.

1.1 QS/QM durch ein kennzifferngestütztes Steuerungssystem

Das System verbindet und integriert zum einen die vielfältigen schon jetzt existierenden Maßnahmen² mit der Erhebung von qualitätsrelevanten Kennziffern (KPI = Key Performance Indicator). Natürlich wird sich zum anderen – im Kontext der angesprochenen Veränderungsprozesse – die Notwendigkeit neuer Maßnahmen und damit verbundener KPIs ergeben. Die KPIs werden in einer Ziel-Matrix »Stakeholder x Bereiche« abgebildet. Als Stakeholder werden alle beteiligten Akteure beginnend bei Schülern (Studien-Interessenten), über Studierende, Absolventen bis hin zum wissenschaftlichen Personal, aber auch Mitarbeitern aus Technik und Verwaltung einbezogen.

Das Qualitätsmanagement-System ermöglicht die Schaffung von Kennzahlen und Steuerung auf Basis dieser Kennzahlen in allen universitären Handlungsfeldern.

Auch in der Forschung muss QM auf Indikatoren zurückgreifen. Indikatoren sind jedoch nur fachspezifisch sinnvoll. Indikatoren bedürfen einer kontinuierlichen Reflexion hinsichtlich ihrer Eignung und ihrer Messeigenschaften – diesen bekannten Forderungen (z.B. des IFQ) kann mit dem QM-System auf der Basis von KPIs entsprochen werden³. Damit wird es möglich, entlang von Entwicklungsstrategien Konsequenzen (incentives und constraints) zu definieren. *Für Qualitätssicherung und Qualitätsmanagement der Forschung – aber nicht nur dort – benötigt man ohne Frage eine solide Informationsgrundlage in Gestalt eines Informationssystems, das über wesentliche Input, Output und Prozessgrößen informiert.*

Dass Evaluation der Lehre und Studienbedingungen natürlich in dem System genau in der beschriebenen Weise eine zentrale Rolle spielen, versteht sich fast von selbst. Schon jetzt wird im Rahmen des Bonner Modells auf der Basis von Indikatoren gearbeitet. Studierende aber auch Absolventen werden über Absolventenbefragungen einbezogen. In den *Zielvereinbarungen III der Universität mit dem Land NRW* und der *Evaluationsordnung* werden die sich aus den Evaluationen ergebenden Konsequenzen thematisiert (z. B. in Form von Zielvereinbarungen, Anreizsystemen oder Sanktionen, Lehrpreisen).

Qualitätssichernde Maßnahmen erstrecken sich auch auf wichtige Schnittstellen, wie Aufnahme des Studiums, insbesondere aber auch Übergang Schule/Hochschule durch aktive Mitwirkung der Hochschule an der Zulas-

2 Qualitätssicherung der Forschung, Kennziffern aus der Personalentwicklung, regelmäßige (Lehr-) Evaluation mit zahlreichen Modulen nach dem Bonner Modell, vgl. Abschnitt 2.2.

3 vgl. dazu besonders *Qualitätssicherung in der Forschung an der Universität Bonn*.

sung zum Studium (Self-Assessment und Studieneingangstests), aber eben auch am Übergang Hochschule/ Beruf durch die erwähnten Absolventenbefragungen mit dem Ziel, eine Absolventenkultur zu etablieren und die Alumni-Vereinigung zu stärken.

Qualitätssicherung durch leistungsgerechte Hochschulfinanzierung wird ermöglicht, da Kennziffern als Steuerungsinformationen für Mittelbemessungsmodelle, Personalbemessungskonzepte, Flächenmanagement, Globalhaushalt, Zielvereinbarungen zur Verfügung stehen.

Über die eigene Hochschule hinaus wird Qualitätsvergleich möglich, Markttransparenz geschaffen. In diesem Sinne fügt sich ein solches System – ganz im Sinne eines umfassenden Change Prozesses – in Professionalisierungsbestrebungen der Hochschulleitung genau so ein wie in neue Organisationsstrukturen (z. B. Hochschulrat) oder neue Rekrutierungsstrategien.

Das KPI-System stellt also für die Qualitätsdimensionen »Konzeptqualität, Strukturqualität, Prozessqualität, Ergebnisqualität und Wirkungsanalyse« die erforderlichen Informationen zur Verfügung und verlangt bei der Definition der Indikatoren »hochschulpolitische« konzeptuelle und strukturelle Entscheidungen.

Aus der Beschreibung des KPI-gestützten Qualitätssicherungs- und -managementsystems wird deutlich, dass die Stakeholder der Hochschule (Studenten, Absolventen, Alumni, Mitarbeiter, Unternehmen etc.) eng ein- und angebunden sind. Gerade Unternehmen sind durch den Zuschnitt des Projekts sehr stark mit der Universität vernetzt. So kooperiert die Universität Bonn (schon jetzt) bei der Entwicklung und Implementierung von Maßnahmen in den Bereichen Qualitätsmanagement und Evaluation u. a. mit dem Institut für Forschungsinformation und Qualitätssicherung (IFQ), IBM, der T-Mobile International sowie der Deutschen Post World Net AG.

1.2 Information und Steuerung

Daten aus bestehenden und neu einzurichtenden Informationsquellen werden automatisiert in das System eingespeist. Ein transparentes Informationssystem mit entsprechendem zentralen Data Warehouse ist die Grundlage des kennzifferngestützten Steuerungssystems (Metadatenbank, s. Abbildung)



Häufig wird der Datenerfassungs- und Auswertungsprozess in seinem Aufwand und in seiner Wichtigkeit für den gesamten Veränderungs- und Verbesserungsprozess unterschätzt. Unabdingbar aber ist, dass Daten im gewünschten Umfang, mit der notwendigen Zuverlässigkeit und Aktualität verfügbar sind, sich in der gewünschten Form auswerten lassen, kohärent und intern konsistent und deren Quellen bekannt sind, auf DV-Systemen gespeichert und nicht verstreut, sondern »zentral« verfügbar sind.

Nur so ist die notwendige (natürlich nicht hinreichende) Voraussetzung gegeben,

- Prioritäten festlegen zu können,
- quantitative Ziele festzulegen,
- Führungskräfte und Mitarbeiter nachhaltig zu überzeugen, dass z. B. die Ist-Situation verbessert werden muss.

Es handelt sich also um kein schieres Controllinginstrument, das (nur) große Mengen an Zahlen sammelt (Monitoring), sondern ein übergeordnetes System, das zwar auf der Erhebung von Zahlen im Sinne von Kennziffern basiert, auf diese also nicht verzichten kann, bei dem aber die Indikatoren eindeutig einem Steuerungszweck zugeordnet werden können.

Es entsteht somit eine effiziente, transparente Steuerungsgrundlage im Sinne eines umfassenden Qualitätsmanagements hinsichtlich Monitoring/Benchmarking, Zielvereinbarungen, Gratifikations- bzw. Belohnungssystemen sowie für ein universitäres Vorschlagswesen auf verschiedenen Aggregationsstufen. Qualitätssteigerung bis hin zur Exzellenz kann demnach angefangen von individuellen Leistungen Studierender oder Wissenschaftler

über aggregierte Kennziffern für Leistungen von Instituten, Zentren, Fakultäten in Forschung, Lehre und Management bis hin zu Förderung der Hochschule durch ihre Absolventen/Alumni durch eine große Bandbreite von Leistungs- und Qualitätsindikatoren auf allen Ebenen abgebildet werden. Das regelmäßige Tracking dieser Indikatoren ermöglicht darüber hinaus eine Beobachtung der Entwicklung über die Zeit, gezielte Anhaltspunkte für Maßnahmen sowie ein professionelles Personalmanagement (via Zielvereinbarungen, Gratifikation, Erfassung des Commitments etc.).

1.3 Organisation und Ressourceneinsatz

Oberste Verantwortlichkeit liegt beim Rektorat, welches einen seiner Prorektoren mit Wahrnehmung der Qualitätssicherung und des Qualitätsmanagements betraut. Gemäß Rektoratsbeschluss vom 7. Mai 2007 ist ein Beauftragter für Qualitätssicherung und Evaluation bestellt worden (Leiter des Zentrums für Evaluation und Methoden – ZEM), um die vielfältigen Ansätze zur Qualitätssicherung und Evaluation an der Rheinischen Friedrich-Wilhelms-Universität Bonn zu *harmonisieren* und vor allem nach *wissenschaftlichen* Standards synergetisch aufeinander abzustimmen und durch *Forschung* zu begleiten (s. a. Abschnitt 2.2).

Analog wird sich auf der Fakultätsebene gemäß den von den Dekanen geforderten Aktivitäten eine funktionelle und strukturelle »Hierarchie« der Verantwortung ergeben. In manchen Fakultäten und Instituten sind die für gute Information, Kommunikation und effiziente Entscheidung hilfreichen und notwendigen Strukturen schon vorhanden. Die erwähnte und natürlich essentielle Steuerungswirkung ist damit nicht auf die Ebene der Hochschulleitung begrenzt, sondern erstreckt sich auf alle Organisationsebenen.

Es wird ein Lenkungsausschuss gebildet, in dem von universitärer Seite der zuständige Prorektor, der Kanzler, der Qualitäts- und Evaluationsbeauftragte, ein Dekan und ein Dezernent vertreten sind.

1.4 Qualitätskultur

Der *IBM Global Business Service* hat in Zusammenarbeit mit dem *Zentrum für Evaluation (ZEM)* der Universität Bonn mehrere Studien vorgelegt: *Making Change Work* (2007, 2008). Wir beziehen uns im Folgenden mehrfach auf diese Studien, nicht nur weil die Universität Bonn an ihr mitgewirkt hat, sondern weil aus ihr auch klare Handlungsanweisungen ableitbar sind⁴.

4 s. im 2. Abschnitt Umsetzung: 2.3 Checkliste für die Projektpraxis, die die Standortbestimmung bezüglich der Professionalität im Change Management erlauben und aufzeigen, wie sich die zentralen Hebel für erfolgreiche Veränderung gezielt nutzen lassen.

Ziel der IBM/ZEM-Studien war es, die wesentlichen Faktoren für die erfolgreiche Umsetzung von Veränderungen auf die Basis empirischer Daten zu stellen. Neben der Bestätigung bekannter Prinzipien – Commitment des Top-Managements und ehrliche Kommunikation anstelle von Schönfärberei – hat sich eine Reihe neuer Erfolgsfaktoren aber auch Hürden herauskristallisiert. Das Verändern von Denkweisen und Einstellungen wird als eindeutig größte Herausforderung identifiziert und liegt damit noch vor den erwartungsgemäß bedeutsamen knappen Ressourcen.

Während Führungskräfte das Scheitern von Veränderungen oft in einer mangelnden Motivation der betroffenen Mitarbeiter suchen, bewerten Praktiker diesen Faktor weit weniger kritisch als das Commitment des Top-Managements zu den getroffenen Entscheidungen. Besonders wichtig für die Praxis sind drei Kernfaktoren, die signifikanten Einfluss auf den Projekterfolg haben.

- Das Bewusstsein für die Herausforderungen, die mit einer Veränderung einhergehen.
- Der Einsatz einer formalen Change-Management-Methodik, welche vor allem integrative und motivierende Kommunikation einschließt (s. im einzelnen Abschnitt 2.1 und 2.3)
- Der Anteil des Projektbudgets, der direkt in Change Management fließt.

Der Universität Bonn ist bewusst, dass »Implementation einer Kultur« ihre Zeit braucht und externe Berater aller Erfahrung nach vonnöten sind. Wie angedeutet, werden aber die vorhandenen Strukturen und Netze für Kommunikation systematisch und regelhaft genutzt werden: Hochschulrat, Rektorat, Senat, Dekanekonferenz, Dekanate, Fakultätsräte, Strukturkommissionen, Geschäftsführer- und Kustodenrunden auf Fakultätsebene, Institutsvorstände, Abteilungssitzungen, Dezernentenrunde, Dezernatsmeetings, AStA, Fachschaftskonferenzen, Fachschaften, »Studiengebühren«-Kommissionen, etc. Es bedarf – zumindest vorerst einmal – keiner neuen Strukturen, einige der existierenden könnten sich gar als überflüssig erweisen. Natürlich müssten u. U. neue hinzukommen, gemäß dem Prinzip *structure follows function*. »Erfolgreiche« Kommunikation (s. Abschnitt 2.3 Checkliste) auf diesen Wegen zeitigt aller Erfahrung nach die Akzeptanz, welche für Veränderung notwendig ist, und der damit verbundenen Maßnahmen (z.B. Einführung eines KPI-gestützten QM-Systems) und führt zu Transparenz von Prozessen und Entscheidungen – unabdingbare Voraussetzungen für die Nachhaltigkeit des Wandels.

2. Umsetzung

Die Etablierung des hochschulweiten kennzifferngestützten Steuerungssystems dient als Basis des Qualitätsmanagements für Forschung, Lehre und

Management. Ziel ist die Einführung einer Steuerungsgrundlage zur Steigerung der Profilbildung und Positionierung der Universität.

2.1 Konkrete Umsetzungsschritte

In einem ersten zentralen Schritt werden qualitätsrelevante Kennziffern (KPIs) aus Hochschulleitbild und Qualitätszielen (s. dazu die in den Fußnoten erwähnten Dokumente) abgeleitet und operationalisiert. Bei der Ableitung von relevanten Indikatoren mit den Beteiligten handelt es sich um einen mühseligen, in jedem Fall extern zu moderierenden Prozess. Dies schließt die Selektion von relevanten Controllingdaten ein, bei gleichzeitiger Berücksichtigung von Leitbild, Profil und Bedürfnissen der verschiedenen Organisationsebenen (wie Hochschulrat, Rektorat, Senat, Fakultäten, Verwaltung, Studierende, Fachschaften). Dazu existieren bislang nur punktuelle Aktivitäten. Deren Bündelung und Integration soll u. a. durch den Qualitäts- und Evaluationsbeauftragten initiiert und unterstützt werden (s.1.3 und 2.2). Erfahrungsgemäß sind diese Kennziffern zum Teil aus bestehenden Datenquellen (z.B. Verwaltungsdaten, Evaluationen, Forschungsberichte etc.) extrahierbar. Für »fehlende«, aber hochschulpolitisch steuerungsrelevante Qualitätsindikatoren sind entsprechende effiziente Erhebungsmodi abzuleiten.

In einem zweiten Schritt werden die KPIs in eine »Metadatenbank« integriert. So entsteht eine effiziente, transparente Steuerungsgrundlage im Sinne eines umfassenden Qualitätsmanagements unter Einbeziehung von Monitoring/Benchmarking, Zielvereinbarungen oder Gratifikations- bzw. Belohnungssystemen auf verschiedenen Organisationsebenen.

Die Implementierung des kennzifferngestützten Steuerungssystems erfolgt in den Schritten: (I) Qualitative Ableitung von Key Performance Indicators (KPIs) und Zustimmung aller beteiligten Akteure zum Verfahren im moderierten Prozess und kontinuierliche Erfassung der Kennziffern, (II) Implementierung des individuell zugeschnittenen Data Warehouse mit den entsprechenden Schnittstellen und ggf. weiterer Tools zur Kennziffernerfassung/-aggregation, (III) Ableitung individueller Reporting-Tools, (IV) Interpretation der Ergebnisse und Maßnahmenableitung, (V) Regelmäßige Überprüfung der Kennziffern und Weiterentwicklung des Systems, (VI) Überprüfung der Wirksamkeit von Maßnahmen.

2.2 Bisheriges Qualitätsmanagement: Elemente und Verantwortlichkeit

An der Rheinischen Friedrich-Wilhelms-Universität Bonn existieren zahlreiche Qualitätssicherungsvorgaben und -maßnahmen, zum einen zentral verankert, wie

- Qualitätskriterien der Universität Bonn
- Qualitätssicherung in der Forschung
- Ziele, Standards und Führungsgrundsätze der Verwaltung

oder die im Rektoratsauftrag vom Zentrum für Evaluation und Methoden (ZEM) durchgeführt und unterstützt werden, wie

- das Bonner Modell der Hochschulevaluation (s. auch Ziel- und Leistungsvereinbarungen – ZLV III 2007 bis 2010 mit dem Land: § 2 (1) Qualitätssicherung in der Lehre, (6) Studienberatung, incl. Self-Assessment und Studieneingangstests (7) Absolventenverbleibs- und -erfolgsstudien vor allem für die Qualitätssicherung des Lehrangebots bzw. des Ausbildungsprofils relevant),
 - das Prüfungsgremium zur Sicherung der Qualität der Lehr- und Studienorganisation (gemäß §12 der Studienbeitrags- und -gebührensatzung der Rheinischen Friedrich-Wilhelms-Universität Bonn),
- zum anderen aber auch dezentral unter Leitung engagierter Fachschaften, Institute und Fakultäten.

Sie betreffen nicht nur Lehre und Studium, sondern – wie die vorgenannten Bereiche zeigen – das gesamte Spektrum des universitären Aktionsraumes, also auch Forschung, Verwaltung, Management.

Die Qualitätskriterien der Universität Bonn beziehen sich auf die Bereiche Forschung: Preise, Forschungsverbünde, Drittmittelwerbung mit wissenschaftlicher Begutachtung, Publikationen/Monographien, Innovationstransfer, Rufe, Habilitationen; Lehre: Stipendien, herausragende Studiengänge, Absolventen, Doktorandenprogramme, Anzahl der Studierenden; Funktion: Vorsitz/Mitgliedschaft in (inter-)nationalen Akademien, Institutionen, Gremien, Leitung von Forschungsverbänden, Berufungen, Herausgeberfunktionen, Einladungen; Organisation: Forschungsförderung, Leistungsorientierte Mittelvergabe, Evaluierung/Controlling, Strukturplanung, Zentraler Servicebereich, Kommunikation und Öffentlichkeitsarbeit. Die Qualitätskriterien der Fakultäten werden in einer Matrix »7 Fakultäten x 4 Bereiche« zum Teil fakultätsspezifisch ausdifferenziert. Das soll den Vergleich zu gleichartigen Fakultäten anderer Universitäten ermöglichen, aber auch zu den anderen Fakultäten innerhalb der Universität. Dies setzt allerdings einen nicht unerheblichen Umfang an gemeinsamen Kriterien für die Fakultäten voraus. An diesen wenigen Beispielen wird einerseits das Potential daraus abgeleiteter KPIs deutlich, aber auch die Problematik, diese konsensual zu konzeptualisieren und zu operationalisieren.

Die Qualitätssicherung in der Forschung verdeutlicht, dass die Universität über ein Bündel von Maßnahmen verfügt, welches auf unterschiedlichen Ebenen wirksam ist, und zwar durch leistungsorientierte Forschungsförderung, parametergesteuerte Unterstützung durch Sachmittel, Anreizsysteme (gestuft in vier Gruppen), universitätsinterne Ausschreibung zur Anschubfinanzierung für wissenschaftliche Gemeinschaftsprojekte, nur um einige Maßnahmen zu nennen (vgl. auch BONFOR der Medizinischen Fakultät). Mit dem Institut für Forschungsinformation und Qualitätssicherung (IFQ) besteht ein Kooperationsvertrag, im Rahmen dessen diese Exzellenz- und

Qualitätskriterien u. U. dergestalt verfeinert und geschärft werden, dass die Vergleichbarkeit auch für den (universitätsinternen) Wettbewerb erhöht wird.

Das Bonner Modell der Hochschulevaluation wurde unter Federführung des Zentrums für Evaluation und Methoden (ZEM) und unter intensiver Beteiligung von Studierenden, Fachschaften und Dozenten entwickelt. Das Modell setzt sich aus folgenden Elementen zusammen:

- Allgemeine Befragung aller Studierenden
- Evaluation von Lehrveranstaltungen
- Zielsetzungen in den Lehrveranstaltungen
- Befragung von Absolventen
- Monitoringdaten aus der Verwaltung
- Panelerhebung
- Studienanfänger- und Hochschulwechslerbefragung

Das Bonner Modell generiert hierzu mittels online-gestützter Befragungsmodule regelmäßig eine breite Datenbasis. Durch die modulare Struktur werden explizit die verschiedenen Perspektiven der Beteiligten sowie die Betrachtung verschiedener Ebenen und Aufgabenbereiche berücksichtigt. Es erfolgt eine Rückmeldung an die Fakultäten und das Rektorat, um vor allem zu gewährleisten: Erstellung von Ziel- und Maßnahmenkatalog unter Beteiligung von Studierenden und Dozenten, Umsetzung der Maßnahme und Erfolgscontrolling durch gezielte Fragen zur Umsetzung der Maßnahmen in den Befragungsmodulen sowie Berichterlegung der Lehreinheiten, Wirkungsanalyse.

Ein besonders interessantes Modul des Bonner Modells ist die online-gestützte Befragung des wissenschaftlichen Personals. Dies bindet die Lehrenden in ein umfassendes Qualitätssicherungskonzept ein. Erfragt werden Informationen über die Arbeitsbedingungen, die eigenen Qualifizierungsbedingungen sowie die Wahrnehmung von Image und Leitbild der Universität. Die Befragung von Mitarbeitern leistet somit einen wichtigen Beitrag für ein professionelles Personalmanagement.

Die *Ziele, Standards und Führungsgrundsätze der Verwaltung* veranschaulichen, wie das Leitbild der Universität zum Fundament und Ausgangspunkt des Handelns gemacht wird. Die Verwaltung lässt sich von folgenden Qualitätsstandards leiten und will an ihnen gemessen werden: *Schnelligkeit, Flexibilität, Wirtschaftlichkeit, Zuverlässigkeit und Planbarkeit, Sachangemessenheit.*

Zur Unterstützung der hieraus resultierenden Leistungs- und Qualitätsprozesse wird zur Zeit ein Projektmanagementsystem eingeführt, das das Konzept der Steuerung über Regelkreisläufe und Rückkopplungseffekte in den Vordergrund stellt. Der im Mai 2007 begonnene Einführungsprozess der Methodik und des dazugehörigen DV-Tools wird parallel durch das ZEM evaluiert mit dem Ziel einer zeitnahen Qualitätssicherung von Einführungsprozess und -ergebnissen. Diese Begleitung hat im Rahmen von Mitarbeiter-

befragungen schon jetzt zu interessanten und verwertbaren Befunden zur Kommunikations- und Qualitätskultur geführt, welche für die Einführung des KPI-Systems sehr hilfreich sein werden.

In den *Führungsgrundsätzen* nehmen sich vor allem die Führungskräfte selbst in die Pflicht, indem folgende Grundsätze mit Leben zu füllen sind: Führungskräfte gewährleisten eine moderne Dienstleistungskultur, sie haben Vorbildfunktion, sie sind kritik- und konfliktfähig, sie treffen Entscheidungen, sie informieren, sie sorgen für Zusammenarbeit, sie führen zielorientiert durch Delegation und Erfolgskontrolle, sie sind für Personalentwicklung verantwortlich⁵

Oberste Verantwortlichkeit liegt beim Rektorat (s. Abschnitt 1.3), welches einen seiner Prorektoren mit Wahrnehmung der Qualitätssicherung und des Qualitätsmanagements betraut. Beim Qualitätsbeauftragten (gleichzeitig Leiter des ZEM) handelt es sich um eine Stabstelle, die die Entwicklung und Einführung koordiniert und vor allem nach Standards synergetisch aufeinander abstimmt.

2.3 Kommunikationskultur

Ans Ziel einer erfolgreichen Implementation eines neuen – wie hier KPI-gestützten – QM-Systems führt nur Kommunikation, die über unterschiedliche interaktive Kanäle und mit einer möglichst großen »Face-to-Face«-Komponente arbeitet. Ein ehrlicher Umgang, bei dem auch »schlechte Nachrichten« frühzeitig und offen kommuniziert werden, erhält und stärkt während der Veränderungsphase das Vertrauen und die Motivation der Mitarbeiter. Die jüngsten Mitarbeiterbefragungen in der Verwaltung der Universität Bonn anlässlich der Einführung eines elektronischen Projektmanagementsystems bestätigen diese Erkenntnis.

Der Umsetzungsprozess wird vor allem im Hinblick auf die Akzeptanz also an folgenden Kriterien orientiert sein, welche in einer Checkliste für die Praxis aus den erwähnten IBM/ZEM-Studien dargelegt sind.

Making Change Work – Checkliste: Eine Checkliste für die Praxis

Begleitende Maßnahmen

- Ist das Commitment des Top-Managements für die Veränderung vorhanden – d.h., ist das Projekt wirklich auf der internen Top-Agenda?
- Wird das Sponsorship vom Top-Management auch vorgelebt (statt an Vertreter, Newsletter oder Frontalveranstaltungen delegiert)?
- Erfolgt die Kommunikation – auch schlechter Nachrichten – rechtzeitig, ehrlich und offen?

5 Vgl. dazu besonders auch den Abschnitt 1.4. und den folgenden Abschnitt 2.3

- Sind persönliche Kommunikation und individuelles Feedback zentrale Bestandteile der Kommunikationsstrategie?
- Ist die informelle Kommunikation Teil der Kommunikationsstrategie, wird sie aktiv unterstützt und für Feedback genutzt?
- Werden die Mitarbeiter aktiv in den Veränderungsprozess eingebunden oder wird eine Einbindung nur nach außen »zelebriert«?
- Sind der Projektverlauf und die Arbeit der Projektgruppe für die Betroffenen ausreichend transparent?
- Werden fachliche Experten eingebunden und bei externen Experten der beidseitige Wissenstransfer sichergestellt?
- Wurde der individuelle Nutzen für jeden Beteiligten formuliert und klar kommuniziert?

Best Practice – Erfolgsfaktoren für die Veränderung

- Existiert in der Organisation ein ausreichendes Bewusstsein für die mit einer Veränderung einhergehende Herausforderung?
- Werden mindestens 10% des Projektbudgets für Change Management eingesetzt?
- Wird eine konkrete, formale Vorgehensweise für Change Management konsistent angewendet?

Vorbereitende Maßnahmen

- Gibt es eine klar kommunizierte Vision und Erwartungen mit konkreten Zielvorgaben?
- Sind die Verantwortungsbereiche für die Veränderung klar definiert und für alle transparent?
- Ist die Komplexität der Projektumgebung berücksichtigt, sind die wichtigsten Abhängigkeiten erfasst und Risiken dokumentiert?
- Ist das Change Management integraler Teil der Projektsteuerung und methodisch eng mit dem Projektmanagement verzahnt?
- Sind zu erwartende kulturelle Widerstände bekannt und eine entsprechende Strategie entwickelt?
- Wurde die Veränderungsfähigkeit und -bereitschaft in der betroffenen Organisationseinheit ermittelt und daraus entsprechende Maßnahmen abgeleitet?
- Wurden Change Agents aus der Organisation identifiziert und proaktiv eingebunden?

Die Checkliste enthält die wesentlichen Facetten erfolgreicher Kommunikation, welche sämtliche Stakeholder einbezieht, und hilft so, Akzeptanz bei allen Akteuren zu erzielen und Transparenz – vor allem auch im Rahmen

moderierter Prozesse – zu gewährleisten⁶. Es geht darum, »Spielregeln«, wenn man so will, integrative und motivierende Kommunikationsrituale im universitären Alltags- und »Geschäfts«leben, zu schaffen, die den Veränderungsbedarf verdeutlichen, die auf entsprechende Auswirkungen vorbereiten, die für gemeinsame Ziele motivieren. Das mag nicht immer ohne Moderation zwischen den verschiedenen Interessengruppen gelingen, muss aber das Commitment der Hochschulleitung, auch des Hochschulrates, sichtbar werden lassen und die Akzeptanz auf Seiten der Führungskräfte für allfällige Veränderungen vertiefen. In Anlehnung an den Bestseller von Spencer Johnson schon aus dem Jahre 1998 mit dem Titel *Who moved my cheese?* kann die »Philosophie« des Bonner Ansatzes folgendermaßen charakterisiert werden: Statt mit den Menschen *Hem & Haw* (Grübel & Knobel) halten wir es mit den Mäusen *Sniff & Curby* (Schnüffel & Wusel) oder statt selbstgefälliger Zufriedenheit und unangemessener Arroganz ist kontinuierliche Suche nach Verbesserung die Triebfeder.

Literaturempfehlungen

- BECKER-Kolle, C., Kraus, G. & Fischer, T. (2006). *Handbuch Change-Management: Steuerung von Veränderungsprozessen in Organisationen. Einflussfaktoren und Beteiligte. Konzepte, Instrumente und Methoden* (2. Aufl.). Berlin: Cornelsen.
- DOPPLER, K. & Lauterburg, C. (2002). *Change Management: Den Unternehmenswandel gestalten* (10. Aufl.). Frankfurt/New York: Campus Verlag.
- DOPPLER, K. (2008). *Der Change Manager: Sich selbst und andere verändern – und trotzdem bleiben, wer man ist*. Frankfurt/New York: Campus Verlag.
- IBM Global Business Services. (2007). *Making Change Work: Erfolgsfaktoren für die Einführung von Innovationen*. Zugriff am 10.09.2009 unter http://www-05.ibm.com/de/pressroom/downloads/mcw_2007.pdf
- IBM Global Business Services. (2008). *Making Change Work: Fortsetzung des Dialogs über das Unternehmen der Zukunft*. Zugriff am 10.09.2009 unter http://www-935.ibm.com/services/de/bcs/pdf/2009/making_change_work.pdf
- JOHNSON, S. (2000). *Die Mäuse-Strategie für Manager: Veränderungen erfolgreich begegnen* (G. Turner, Übers.). München: Ariston. (Original erschienen 1998: *Who moved my cheese?*)
- KEUPER, F. & Groten, H. (Hrsg.). (2007). *Nachhaltiges Change Management. Interdisziplinäre Fallbeispiele und Perspektiven*. Berlin: Springer.
- KOGOJ, T. & Mitschka, K. (2004). *Wandelwörter. Konversationslexikon zu Change and Management*. Wien: Linde.
- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2009). *Leitbild der Universität / Our vision* [Broschüre]. Zugriff am 10.09.2009 unter http://www3.uni-bonn.de/die-universitaet/publikationen/ubo_leitbild_09_ansicht_neu.pdf

6 Ansonsten sei auf die Passagen im Abschnitt 1.4 Qualitätskultur verwiesen.

- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2006). *Qualitätssicherung in der Forschung* [Broschüre]. Zugriff am 10.09.2009 unter http://www3.uni-bonn.de/die-universitaet/publikationen/qs_web.pdf
- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2007). *Qualitätskriterien der Universität* [Broschüre]. Zugriff am 10.09.2009 unter http://www3.uni-bonn.de/die-universitaet/publikationen/Qua_kriterien_web_062006.pdf
- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2009). *Verwaltung der Universität Bonn: Ziele und Führungsgrundsätze* [Broschüre]. Zugriff am 10.09.2009 unter http://www3.uni-bonn.de/die-universitaet/publikationen/ubo_verw._zsf_web.pdf
- STOLZENBERG, K. & Heberle, K. (2006). *Change Management: Veränderungsprozesse erfolgreich gestalten – Mitarbeiter mobilisieren*. Berlin: Springer.

Katharina Hörsch & Georg Rudinger

Qualitätssicherung von Lehre und Studium an Hochschulen: Von Evaluationsergebnissen zu Maßnahmen – aktuelle Entwicklungen an der Universität Bonn

Festgelegt durch das Hochschulrahmengesetz und die Hochschulgesetze der Länder, wird Qualitätssicherung und –management zu einem verbindlichen Auftrag für Hochschulen (Wissenschaftsrat, 2006). »Evaluation ist untrennbar mit Qualitätssicherung und –entwicklung verbunden« (Böttcher, Holtappels & Brohm, 2006, S. 7). An Hochschulen setzt sich die Qualitätssicherung von Lehre und Studium aus Ziel- und Leistungsvereinbarungen der Hochschulen mit den Ländern und der Evaluation von universitären Aufgaben zusammen, wobei letztere häufig auf Qualitätsbewertung der Lehre durch Studierende basiert. Zudem gewinnt Evaluation im Zuge der Strukturreform durch die Einführung von Bachelor- und Masterstudiengängen, wachsender Hochschulautonomie und zugleich verstärkten Wettbewerbs der Hochschulen untereinander als Bestandteil der Qualitätssicherung von Studium und Lehre, aber auch der Forschung zunehmend an Bedeutung (Meyer-Guckel & Daniel, 2004). »Die Universitäten haben die Freiheit, sich diejenigen Studierenden auszuwählen, die ihre Anforderungen am besten erfüllen und die Reputation ihrer Hochschule steigern« (Frey, 2006). Wer langfristig die besten Studierenden an sich binden und sich als attraktiver Studienstandort etablieren möchte, der wird die Bedeutung der Qualität von Lehre und Studium angemessen hoch einzuschätzen haben.

Fortlaufende Qualitätssicherung durch Evaluation erfordert die Einbeziehung aller beteiligten Akteure auf verschiedenen Ebenen (Studierende, Lehrende, Absolventen), um zu umfassender Information und Analyse von Stärken und Schwächen aus verschiedenen Perspektiven zu gelangen. Ziel ist dabei die Etablierung eines Qualitätszyklus (vgl. Abbildung 1). Dieser sollte aus der Evaluation von Studium und Lehre, Ableitung von Maßnahmen aus den Evaluationsergebnissen und der Umsetzung sowie Überprüfung dieser Maßnahmen bestehen. Die Maßnahmen sollten sich an den Charakteristika der Hochschule orientieren, um das Profil der Hochschule im Wettbewerb zu stärken. Zur Erfassung der notwendigen prozess- und ergebnisrelevanten Daten müssen Qualitätssicherungsverfahren in die strategische Ausrichtung integriert und an das Profil einer Hochschule angepasst werden. Effizienz,

Nachhaltigkeit und Ressourcenschonung sollten bei der Einführung und Anwendung von Qualitätssicherungsverfahren im Vordergrund stehen.

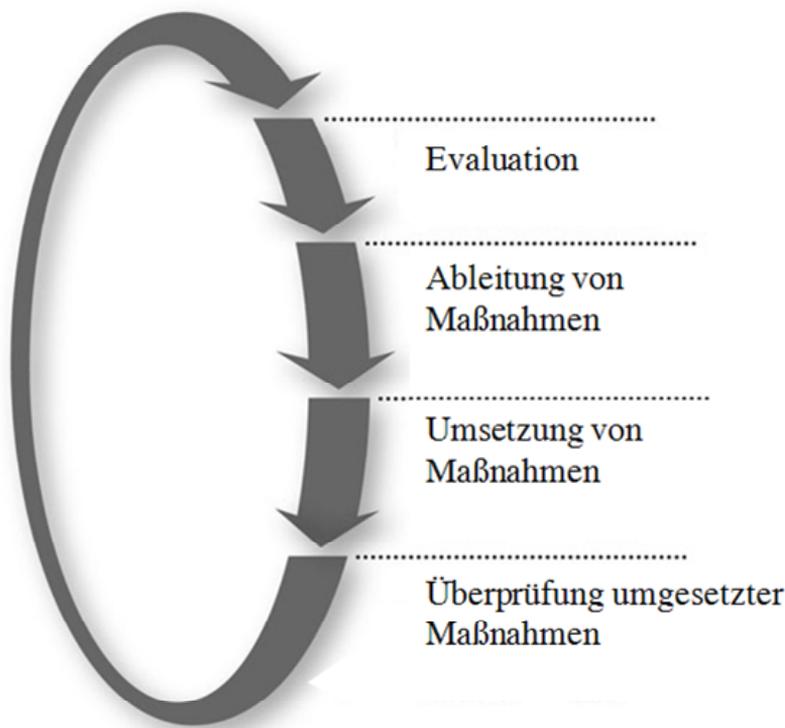


Abbildung 1: Qualitätszyklus

Qualitätssicherung durch Evaluation am Beispiel des Bonner Modells der Hochschulevaluation

Das Bonner Modell der Hochschulevaluation ist als onlinegestütztes Instrument zur Qualitätssicherung konzipiert worden, welches entsprechend der oben dargestellten Anforderungen eine umfassende und zugleich ökonomische Datenerfassung und -auswertung gewährleistet (Rudinger, Krahn, Rietz, 2007). Dem Bonner Modell liegt eine professionelle Software für Umfrageforschung zugrunde, mit der regelmäßig eine breite, empirische Datenbasis gewonnen wird. Durch flächendeckende Erhebungen bei Studierenden, Lehrenden sowie Absolventinnen und Absolventen werden sämtliche Beteiligte einbezogen. Dadurch werden Aspekte, die zur Qualität von Studium und Lehre beitragen, aus verschiedenen Perspektiven bewertet. Somit entsteht eine systematische, empirisch fundierte Grundlage, die eine

tragfähige Interpretation und solide Ableitung von Maßnahmen auf der Ebene von Studienfächern, Fakultäten oder der gesamten Universität ermöglicht.

Ein Studiengang muss sich nach der Akkreditierung und in Abgrenzung dazu in seiner Konzeption und Qualität zunächst bewähren. Mit der online-gestützten Evaluation nach dem Bonner Modell kann aufgrund einer breiten empirischen Basis die Bewertung der notwendigen Qualitätsmerkmale eines Studiengangs z.B. für die Reakkreditierung erfolgen.

Konzipiert und durchgeführt werden die Befragungen vom Zentrum für Evaluation und Methoden (ZEM) im Auftrag des Rektorats in enger Abstimmung mit Vertretern aus Fakultäten und Instituten. In dieser Rolle kann das ZEM die Evaluation standardisiert und automatisiert einerseits für die Universität insgesamt, andererseits auch abgestimmt auf die Erfordernisse einzelner Fächer oder Fakultäten ausrichten. Dementsprechend ist eine Bewertung der zentralen Aufgaben der Universität (u.a. Befähigung zum wissenschaftlichen Arbeiten und zur Übernahme von Verantwortung, Berufsausbildung) sowie der universitären Rahmenbedingungen und -angebote in die Befragungsmodule integriert (vgl. Heidemann, Rietz, Krahn, Riek & Rudinger, 2005).

Die verschiedenen Befragungsmodule des Bonner Modells ermöglichen weitgehend Totalerhebungen, so dass umfassende und personengruppen-spezifische Informationen gewonnen werden können. Des Weiteren verbindet das Bonner Modell die verschiedenen Perspektiven der Beteiligten mit der Betrachtung verschiedener Ebenen und Aufgabenbereiche. Demnach werden Lehre und Studium, Forschung, Nachwuchsförderung und Chancengleichheit aus drei Perspektiven betrachtet: Studierenden- und Studienanfängerbefragung, Absolventenbefragung und die Befragung des wissenschaftlichen Personals werden als onlinegestützte Totalerhebungen durchgeführt. Dies ermöglicht die Gewinnung von Erkenntnissen über die Innen- und Außenwahrnehmung der Universität und ihrer Fakultäten sowie über die Fortschritte bei der Etablierung eines einheitlichen Leitbildes bzw. einer »Organisationskultur«. Lehrveranstaltungen und Module werden durch ein weiteres Befragungsmodule onlinegestützt evaluiert. Dabei steht neben der Bewertung von Didaktik und Vermittlung von Inhalten durch den Dozenten die Erreichung der in den Modulhandbüchern dargelegten Ziele im Vordergrund. Die Flexibilität des Befragungssystems ermöglicht neben dem Einsatz des Standardfragebogens die Anpassung und Ergänzung studiengang- und dozenten-spezifischer Fragen. Als Onlineverfahren ermöglicht es eine automatisierte und effiziente Datenauswertung. Den Dozierenden/Studierenden kann so schnell ein differenziertes Feedback gegeben werden. Die Fakultäten und das Rektorat erhalten ebenfalls eine Ergebnisrückmeldung auf entsprechendem Aggregationsniveau, so dass der Qualitätssicherungs- und -kontrollprozess von verantwortlicher Stelle angestoßen werden kann. Die

dafür notwendige Rücklaufquote ist allerdings auch vom Engagement der Lehrenden und Studierenden abhängig.

Die Absolventenbefragung bietet die Möglichkeit, den Erfolg und Verbleib von Absolventen der Universität Bonn zu verfolgen, Stärken und Schwächen des Studiums an der Universität Bonn rückblickend bewerten zu lassen und dadurch Profil und Leitbild der Universität zu schärfen. Damit ist die Absolventenbefragung Bestandteil eines umfassenden Qualitätssicherungs- und Steuerungssystems, das hier auch eine Brückenfunktion zwischen Universität und Beruf übernimmt. Ziel ist es, eine Absolventenkultur zu etablieren und dadurch den langfristigen Erfolg der Absolventen nachzuvollziehen.

Insgesamt wird mit dem Bonner Modell der Ansatz verfolgt, qualitätssichernde Maßnahmen für Lehre und Studium auf alle wichtigen Schnittstellen einer akademischen Ausbildung zu beziehen. Im Fokus stehen dabei neben der Aufnahme des Studiums insbesondere der Übergang Schule/Hochschule durch aktive Mitwirkung der Hochschule an der Zulassung zum Studium (z.B. Studieneingangstests) und der Übergang Hochschule/Beruf durch die Absolventenbefragung.

Die Integration der Evaluationsergebnisse in den Qualitätssicherungszyklus erfordert einen strukturellen Rahmen, der die Ableitung von Maßnahmen, ihre Umsetzung und die Überprüfung der umgesetzten Maßnahmen regelt. Nur so können Erfolg und Nutzen von Veränderungen bewertet und nachhaltig überprüft werden.

Umsetzung von Evaluationsergebnissen an der Universität Bonn

Bislang gab es an der Universität Bonn keine systematische Ableitung von Maßnahmen, da in erster Linie der Anreiz und die Integration in ein Konzept fehlten. Im August 2008 hat die Universität Bonn eine Evaluationsordnung für Lehre und Studium (Universität Bonn, 2008) mit einer Laufzeit von vier Jahren verabschiedet, die die Rahmenbedingungen der Evaluation formuliert. Damit werden Insellösungen einzelner Institute hinfällig und die Ableitung und Umsetzung von Maßnahmen zur Verbesserung von Studium und Lehre auf verschiedenen Ebenen (Lehreinheiten, Fakultäten, Rektorat) verpflichtend.

Vorgesehen ist die Gründung einer Evaluationsprojektgruppe, bestehend aus einem Evaluationsbeauftragten, Professoren, wissenschaftlichen Mitarbeitern sowie Studierenden auf der Ebene einer Lehreinheit. Zur Ableitung gezielter Maßnahmen aus den Ergebnissen insbesondere der Modul- und Lehreevaluation wird eine ausführliche Diskussion mit den Studierenden empfohlen, die besonders in Verbindung mit der Verwendung von Studienbeiträgen einer Bedarfsanalyse gleichkommt. Ergebnisse der Lehr- und Modulevaluation können beispielsweise auf Lehreinheits- oder Fakultätsbe-

ne Auswirkungen auf Lehrangebot, Ausstattung oder strukturelle Gegebenheiten haben. Für und durch Evaluation und ihre Konsequenzen werden regelmäßige und verbindliche Kommunikationsprozesse initiiert und etabliert. Jedoch braucht die »Implementation« einer solchen Kultur ihre Zeit.

Auf Basis der erhobenen Evaluationsdaten und zentral erfasster Monitoringdaten werden auf Ebene der Lehreinheiten Ziel- und Maßnahmenkataloge entwickelt, die wiederum auf Fakultätsebene gebündelt und vom Dekan zu einem für die Fakultät gesamtgültigen Ziel- und Maßnahmenkatalog zusammengestellt werden. Diese »Fakultätskataloge« erhält das Rektorat und entscheidet (unter Berücksichtigung der Haushaltslage) über die konkrete Umsetzung von Maßnahmen.

Empfehlenswert in diesem Kontext ist die Veröffentlichung von Evaluationsergebnissen in aggregierter Form, um die Akzeptanz des gesamten Verfahrens und die Nachvollziehbarkeit von Maßnahmen zu gewährleisten. Dies trägt genauso zur Verbindlichkeit und Verstetigung des Evaluationsprozesses bei wie Erfolgscontrolling durch gezielte Fragen zur Umsetzung von Maßnahmen sowie Berichterlegung der Lehreinheiten und Wirkungsanalysen. Sonst bleibt Evaluation ein »Glasperlenspiel«.

Gemäß der Evaluationsordnung (LEvO) ist ein universitätsinterner Evaluationsprozess vorgesehen, der sich aus dem Standardverfahren der Datenerhebung – bestehend aus Modulevaluation und Lehrveranstaltungskritik jedes Semester, allgemeiner Studierendenbefragung jährlich sowie Absolventenbefragung ein Jahr, fünf und zehn Jahre nach Abschluss des Studiums – und der u.a. aus diesen Daten resultierenden Konsequenzen und Maßnahmen zusammensetzt (vgl. Abbildung 2). Ziel dieses Prozesses ist die Identifikation von Stärken und Schwächen in Bezug auf Studium und Lehre an der Universität Bonn.

Standardverfahren zur online-gestützten Datenerhebung (Bonner Modell)		
Befragungsmodule (gemäß § 4 Abs. 2 LEVO)*		
WS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	Allgemeine Studierenden-befragung (Ende des WS)
SS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	
WS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	Allgemeine Studierenden-befragung (Ende des WS)
SS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	
WS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	Allgemeine Studierenden-befragung (Ende des WS)
SS	Qualitätssicherung - Modulevaluation (am Ende des Moduls) - Lehrveranstaltungskritik (Zeitpunkt frei wählbar)	

Absolventenbefragung

individuell nach Abschlusszeitpunkt (1, 5 und 10 Jahre)

Abbildung 2: Universitätsinterner Evaluationsprozess gemäß Evaluationsordnung der Universität Bonn

Die onlinegestützte Datenerhebung wird durch Daten aus der zentralen Universitätsverwaltung und den Fakultäten zur Erstellung der Ziel- und Maßnahmenkataloge ergänzt. Auf dieser Basis werden Evaluationsberichte der Fakultäten und des Rektorats erstellt (vgl. Abbildung 3).

3-Jahreszyklus der Evaluation von Studium und Lehre	
(gemäß §§ 5,6 LEvO)	
<p>Bereitstellung der Datenbasis:</p> <p>a) Standardverfahren (Bonner Modell) b) Daten der zentralen Universitätsverwaltung c) Daten der Fakultäten</p> <div style="border: 1px solid black; padding: 10px; margin: 10px auto; width: 80%;"> <p style="text-align: center;">Ziel- und Maßnahmenkatalog der Fakultäten (max. 5 Seiten) (strategische und größere Maßnahmen)</p> <p style="text-align: center;">Evaluationsberichte der Fakultäten (max. 5 Seiten)</p> <p style="text-align: center;">Evaluationsbericht des Rektorats</p> </div>	1. Jahr
	2. Jahr
	3. Jahr

Abbildung 3: Berichtszyklus gemäß Evaluationsordnung der Universität Bonn

Modelle für die erfolgreiche Ableitung von Maßnahmen aus der Evaluation zeigen sich beispielsweise am Institut für Psychologie. Hier wird bereits seit einigen Jahren der »Goldene Geist« als Preis für die beste Lehre, gegliedert nach Veranstaltungstypen, durch die Fachschaft verliehen. Zudem wurde das Instrument Modulkonferenz etabliert. Dabei handelt es sich um die aktive Auseinandersetzung mit den Evaluationsergebnissen und Erfahrungen mit Modulen und Lehrveranstaltungen. Aufgabe der Modulkonferenz ist, die Anpassung an erforderliche Gegebenheiten zu fördern und Umstrukturierungsprozesse zu initiieren.

In der Fachgruppe Biologie werden auf Initiative der Fachschaft alle Evaluationsergebnisse veröffentlicht. Die dadurch erhöhte Transparenz der Lehrveranstaltungsbeurteilung fördert insbesondere die Kommunikation zwischen Studierenden und Lehrenden und trägt so dazu bei, dass sich künftige Lehre an den Bedürfnissen der Studierenden orientieren kann.

Die Universität Bonn hat es sich zum Ziel gesetzt, Qualitätssicherung in ein System unter Berücksichtigung aller Akteure zu integrieren. Angefangen bei der Lehrveranstaltungs- und Modulbewertung über allgemeine Studierendenbefragung zu Absolventenbefragung gilt es, die Ergebnisse aller Befragungen heranzuziehen, um somit Maßnahmen auf allen erforderlichen Ebenen ableiten, umsetzen und ihre Umsetzung evaluieren zu können. Notwendig dafür ist die »differenzierte Herangehensweise an das Evaluationsvorhaben« (Kromrey, 2003), um die »richtigen« Maßnahmen aus den Ergebnissen ableiten zu können.

Dabei obliegt die Sicherung der Ableitung von Maßnahmen und Konsequenzen der Hochschulleitung. Die Veröffentlichung des Evaluationsberichtes der Universität Bonn dient der Präsentation positiver Wirkungen von Maßnahmen zur Qualitätssicherung. Das in Abbildung 2 dargestellte interne Berichtssystem entspricht neben den Evaluationsanforderungen den Kriterien für die Systemakkreditierung (Akkreditierungsrat, 2008).

Als Instrument zur Qualitätssicherung ist Evaluation nur dann dienlich, wenn den Ergebnissen Konsequenzen folgen. Diese Konsequenzen sollten Anreize zur Beseitigung von Schwächen schaffen, beispielsweise durch zusätzliche finanzielle Unterstützung.

Im Vergleich zu den vielfältigen Maßnahmen an anderen Universitäten, wie Weiterbildungsangebote für Lehrende, Wettbewerbe, Sanktionssysteme auf der einen oder Lehrpreise auf der anderen Seite, fokussiert Qualitätsverbesserung an der Universität Bonn Rückkoppelungseffekte aus verschiedenen Befragungen. Immer mehr Lehrende begreifen – nicht erst aufgrund der Evaluationsordnung – Evaluation als Chance für die Lehre und gute Evaluationsergebnisse als Anerkennung für Qualität. Evaluation gibt hier den Anstoß für eine zunehmende Sensibilisierung. Zugleich haben die Studierenden ein Mitsprache- und Steuerungsinstrument an der Hand, das es Ihnen ermöglicht Einfluss auf ihre individuelle Studiensituation zu nehmen. Die Qualität der Lehre an der Universität Bonn hat sich durch Qualitätssicherungsansätze bereits verändert und wird sich durch die regelmäßige Evaluation und ihre Konsequenzen kontinuierlich verbessern.

»Bei realistischer Betrachtung muss man davon ausgehen, dass die Aussagen einer Evaluationsstudie im Bildungssektor nicht den Sicherheitsgrad von Aussagen der psychologischen Grundlagenforschung erreichen können. Dies ist aber auch nicht das Ziel solcher Studien. Die Evaluation hat immer dann ihren Zweck erfüllt, wenn sie die Diskussion um die »beste« Maßnahme oder die »optimale« Gestaltung versachlicht und damit auf eine rationalere Grundlage gestellt hat, auf der schließlich auch ein angemessenes Qualitätskonzept aufgebaut werden kann« (Wottawa, 2006, S. 686) .

Insgesamt ist die Universität Bonn auf einem guten Weg, Evaluationsergebnisse für die gesamte Universität gezielt für die Analyse von Stärken und Schwächen und die Ableitung konkreter Handlungen zu nutzen. Das Ziehen

von Konsequenzen aus der Evaluation zur nachhaltigen Verbesserung der Qualität von Studium und Lehre ist als Teil einer effizienten und transparenten Steuerungsgrundlage für umfassendes Qualitätsmanagement zu sehen, welches die strategische, technologische, strukturelle und personelle Dimension der Universität einbezieht.

Literatur

- AKKREDITIERUNGSRAT. (2008). *Kriterien für die Systemakkreditierung* (Drs. AR 67/2008). Zugriff am 24.08.2009 unter http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/Kriterien_Systemakkreditierung_31_10_08.pdf
- BÖTTCHER, W., Holtappels, H.G. & Brohm, M. (Hrsg.). (2006). *Evaluation im Bildungswesen*. Weinheim, München: Juventa.
- FREY, B.S. (2006). *Evaluitis – eine neue Krankheit* (Working Paper Series). Zürich: Institute for Empirical Research in Economics, University of Zurich.
- HEIDEMANN, K., Rietz, C., Krahn, B., Riek, S. & Rudinger, G. (2005). Das Bonner Modell der Hochschulevaluation: Ein universitätsweites Verfahren zur Bewertung von Hochschulleistungen. *Zeitschrift für Evaluation*, 2, 195-219.
- KROMREY, H. (2003). Qualität und Evaluation im System Hochschule. In R. Stockmann (Hrsg.), *Evaluationsforschung* (2. Auflage, S. 233-258). Opladen: Leske+Budrich. Zugriff am 02.09.2008 unter http://www.profkromrey.de/Kromrey_Eval_Univ.pdf
- MEYER-Guckel, V. & Daniel, H.-D. (2004). Bewährung und Wirksamkeit der Verfahren der zentralen Evaluations- und Akkreditierungsagentur Hannover (ZEvA) und des Verbundes Norddeutscher Universitäten (VNU) für die Evaluation von Studium und Lehre. In HRK (Hrsg.), *Metaevaluation. Evaluation von Studium und Lehre auf dem Prüfstand* (Beiträge zur Hochschulpolitik 5/2004, S. 17-28). Zugriff am 28.08.2009 unter <http://www.hrk.de/de/download/dateien/Metaevaluation.pdf>
- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2008). Neufassung der Evaluationsordnung für Lehre und Studium der Rheinischen Friedrich-Wilhelms-Universität Bonn (LEvO) vom 4 August 2008 (Amtliche Bekanntmachungen, 38. Jahrgang, Nr. 26 12). Zugriff am 28.08.2009 unter <http://hss.ulb.uni-bonn.de/amtliche-informationen/aml.bekanntmachungen/2008/26.pdf>
- RUDINGER, G., Krahn, B. & Rietz, C. (2007). Möglichkeiten und Perspektiven online-gestützter Verfahren im Kontext von nachhaltiger Qualitätssicherung und Evaluation. In R. Keil, M. Kerres & R. Schulmeister (Hrsg.), *e-University – Update Bologna* (S.223-242). Münster: Waxmann.
- WISSENSCHAFTSRAT. (2006). *Empfehlungen zur künftigen Rolle der Universitäten im Wissenschaftssystem* (Drs. 7067-06). Zugriff am 28.08.2009 unter <http://www.wissenschaftsrat.de/texte/7067-06.pdf>
- WOTTAWA, H. (2006). Evaluation. In A. Krapp & B. Weidenmann, *Pädagogische Psychologie: Ein Lehrbuch* (S. 659-685). Weinheim: Beltz PVU.

René Costa

Das Bonner Modell der Hochschulevaluation: Erste Erfahrungen mit der Implementierung einer webbasierten Modul- und Lehrveranstaltungsevaluation

Seit nunmehr einem Jahr führt das Zentrum für Evaluation und Methoden (ZEM) als zentrale Einrichtung unter der Verantwortung des Rektorats der Rheinischen Friedrich-Wilhelms-Universität Bonn eine zentrale Modul- und Lehrveranstaltungsevaluation in den neuen modularisierten Studiengängen durch. Dabei kommt das Bonner Modell der Hochschulevaluation zur Anwendung. Im Folgenden werden die ersten – im Zusammenhang mit der Umsetzung der Modul- und Lehrveranstaltungsevaluation erworbenen – Erfahrungen beschrieben.

Als Rechtsgrundlage für die Evaluation dient die Neufassung der Evaluationsordnung für Lehre und Studium der Rheinischen Friedrich-Wilhelms-Universität Bonn (LEvO) vom 4. August 2008 (vgl. hierzu auch Hörsch & Rudinger zur Qualitätssicherung von Lehre und Studium an Hochschulen im selben Band). Insbesondere drei Aspekte der LEvO haben sich als sehr bedeutsam für die Umsetzung erwiesen:

1. Das Bonner Modell der Hochschulevaluation des ZEM wird als Regelverfahren für die Evaluation von Studium und Lehre in den neuen modularisierten Studiengängen an der Universität Bonn eingesetzt.
2. Für jede Lehreinheit benennt der zuständige Dekan einen Evaluationsbeauftragten. Diesem steht eine auf Lehreinheitsebene gegründete Evaluationsprojektgruppe bestehend aus Professoren, Wissenschaftlichen Mitarbeitern und Studenten zur Seite.
3. Die regelmäßige Evaluation dient der Ableitung und Realisierung von Maßnahmen zur Qualitätssicherung und -entwicklung.

Zu Beginn der Umsetzung der Evaluationsordnung stand die Entwicklung eines neuen Befragungsinstrumentes zur Modul- und Lehrveranstaltungsevaluation. Ausgangsbasis dafür war ein bereits seit mehreren Jahren vom ZEM verwandter Fragebogen. Dieser kam innerhalb eines webbasierten Onlinebefragungssystems für die Evaluierung von Lehrveranstaltungen in den »alten« nicht-modularisierten Studiengängen zum Einsatz. Der vorhandene lehrveranstaltungsbezogene Fragebogen wurde einer statistischen und inhaltlichen Analyse und Überarbeitung unterzogen. Dabei wurden auch Rückmeldungen der Studenten zum vorhandenen Fragebogen einbezogen.

Die Ziele der Überarbeitung waren, den neuen Fragebogen möglichst kurz zu fassen und die Ableitung von Maßnahmen aus dem zu erhaltenden Datenmaterial zu optimieren. Ferner wurde der Fragebogen um einen modulbezogenen Fragenteil erweitert, der den speziellen Anforderungen modularisierter Studiengänge Rechnung trägt. Der Fragebogen wurde insgesamt so konstruiert, dass er im Wesentlichen für die meisten Studiengänge geeignet sein sollte.

Um Vergleiche auch über verschiedene Studiengänge zu ermöglichen, wird der Fragenkatalog des modulbezogenen Teils als feste Vorgabe für alle Studiengänge angesehen. Er kann nicht gekürzt werden, es sind jedoch Ergänzungen auf Fakultäts- oder Lehreinheitsebene möglich. Unter Lehreinheit wird dabei eine (oder mehrere gemeinsam) einen Studiengang anbietende Organisationseinheit(en) verstanden. Im Lehrveranstaltungsbezogenen Fragenteil sind Kürzungen oder Modifikationen möglich, sofern diese für die Lehreinheit wichtig und erforderlich sind. Hier können auch spezifische Fragen auf Fakultäts- oder Lehreinheitsebene ergänzt werden, wie beispielsweise nach der Verwendung von Studienbeiträgen, und vor allem auch solche nach dem Erfolg der mit Studienbeiträgen finanzierten Maßnahmen. Darüber hinaus können im Lehrveranstaltungsbezogenen Fragenteil von den jeweiligen Lehrenden eigene Fragen für Ihre Lehrveranstaltungen ergänzt werden.

Entsprechend dem Bonner Modell der Hochschuleevaluation wird auch die Modul- und Lehrveranstaltungsevaluation mit einem Onlinebefragungssystem umgesetzt. Denn eine flächendeckende Evaluation ist an einer Volluniversität wie der Universität Bonn mittels Paper-Pencil-Methode nur mit erheblichem personellen und zeitlichen Aufwand zu bewältigen (Heidemann, Rietz, Krahn, Riek & Rudinger, 2005). Das ZEM verwendet für sein onlinegestütztes Befragungssystem softwareseitig eine Kombination aus der Befragungssoftware SPSS mrInterview, dem Microsoft SQL Server als Datenbank-Server sowie der Programmiersprache Python.

Die Evaluationsordnung der Universität Bonn regelt die Zuständigkeiten und die Aufgabenverteilung auf den verschiedenen Ebenen der universitären Organisationsstruktur. Generell ist das Rektorat für die Durchführung der Evaluation verantwortlich. Innerhalb der Fakultäten sind die jeweiligen Dekane für die Hochschuleevaluation zuständig. Auf der Ebene der Lehreinheiten koordinieren die vom Dekan ernannten Evaluationsbeauftragten das Verfahren.

Für das ZEM fungieren die Evaluationsbeauftragten als zentrale Ansprechpartner in den Lehreinheiten, mit denen beispielsweise erforderliche Anpassungen des Fragebogens abgestimmt werden. Üblicherweise regeln sie auch die Erfassung der für die Modul- und Lehrveranstaltungsevaluation relevanten Informationen über Module und Lehrveranstaltungen (Namen der Module und Lehrveranstaltungen, Namen der Lehrenden, Befragungszeiträume etc.) im Onlinebefragungssystem. Der dafür erforderliche Aufwand

soll zukünftig durch einen regelmäßigen Datentransfer aus dem elektronischen Vorlesungsverzeichnis und dem zentralen Prüfungsverwaltungssystem der Universität Bonn in das Onlinebefragungssystem deutlich reduziert werden. Aus Datenschutzgründen würde dieser Transfer ausschließlich Strukturdaten umfassen, wie zum Beispiel die Namen und Nummern der Lehrveranstaltungen und Module. Nach der Durchführung einer entsprechenden IT-Sicherheitsüberprüfung und der Verabschiedung der (die LEvO spezifizierenden) Verfahrensrichtlinien durch das Rektorat kann die Implementierung der zum Datentransfer benötigten Schnittstellen erfolgen. Zu diesem Zweck müssten allerdings weitere finanzielle Mittel zur Verfügung gestellt werden.

Ferner sind die Evaluationsbeauftragten gemäß LEvO für die Ableitung von Maßnahmen zur Verbesserung der Qualität der Lehre in ihrer Lehreinheit zuständig.

Die bisherigen Erfahrungen an der Universität Bonn haben gezeigt, dass für die Einführung der Modul- und Lehrveranstaltungsevaluation in einer Lehreinheit ein Zeitraum von ungefähr vier Monaten benötigt wird. Der Einführungsprozess ist in der Regel in folgende Bestandteile gegliedert:

- Es erfolgt eine Kontaktaufnahme der Lehreinheit mit dem ZEM.
- In einer Vorbesprechung mit dem Evaluationsbeauftragten und der Evaluationsprojektgruppe der Lehreinheit erläutert das ZEM das Bonner Modell der Hochschulevaluation und den Fragebogen.
- Der ZEM-Standardfragebogen wird innerhalb der Evaluationsprojektgruppe diskutiert, und es werden Anpassungswünsche formuliert.
- Die Änderungswünsche am Fragebogen werden mit dem ZEM abgestimmt.
- Das ZEM sorgt für die Einbindung des lehrheitsspezifischen Fragebogens in das Onlinebefragungssystem.
- Unter der Koordination des Evaluationsbeauftragten erfolgt die Erfassung der zur Evaluation erforderlichen Informationen (Modul- und Lehrveranstaltungsnamen, Befragungszeiträume etc.) im Onlinebefragungssystem.
- Vor Beginn der Befragungszeiträume unterrichtet der Evaluationsbeauftragte die betroffenen Lehrenden über das Evaluationsverfahren.

Nach Abschluss der Befragungen erfolgt entsprechend der LEvO die Übermittlung der Befragungsergebnisse an den Evaluationsbeauftragten, der die Ergebnisse zusammen mit der Evaluationsprojektgruppe als Grundlage für die Ableitung von Maßnahmen zur Qualitätsentwicklung verwendet.

Durch die Etablierung einer solchen Evaluationskultur an der Universität Bonn werden jedoch auch Ängste auf Seiten der Lehrenden geweckt.

Als Sorge über eine zentral durch das ZEM durchgeführte Modul- und Lehrveranstaltungsevaluation wurden beispielsweise datenschutzrechtliche Bedenken formuliert. Das ZEM verfügt zum einen über langjährige Expertise auf dem Gebiet der Markt- und Sozialforschung. Dazu gehört die Einhaltung

des Bundesdatenschutzgesetzes (BDSG) sowie darüber hinaus der »Standards zur Qualitätssicherung in der Markt- und Sozialforschung« des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM), der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Bundesverbandes Deutscher Markt- und Sozialforschung (BVM). Aspekte des Datenschutzes fanden bereits bei der Konzeption der LEvO besondere Berücksichtigung und wurden dort ausführlich geregelt. Die Umsetzung der Modul- und Lehrveranstaltungsevaluation durch das ZEM erfolgt unter strenger Einhaltung dieser Vorgaben. Zum anderen finden alle ZEM-Evaluationsaktivitäten in enger Abstimmung mit dem Justizariat und dem Datenschutzbeauftragten der Universität Bonn statt. Schließlich wird durch die bereits erwähnte IT-Sicherheitsüberprüfung und die vom Rektorat zu verabschiedenden Verfahrensrichtlinien etwaigen Bedenken sowohl bezüglich potentieller Hard- und Softwarerisiken als auch bezüglich hochschulpolitischer Aspekte Rechnung getragen. Die angesprochenen Verfahrensrichtlinien regeln unter anderem, welche Instanz welche Evaluationsergebnisse auf welchem Aggregationsniveau zur Verfügung gestellt bekommt.

Einer der genannten hochschulpolitischen Aspekte besteht in der Befürchtung einer möglichen Einflussnahme der Ergebnisse der Modul- und Lehrveranstaltungsevaluation auf Bleibeverhandlungen. Die Einsichtnahme des Rektors in das Datenmaterial zu diesem Zweck ist grundsätzlich nicht auszuschließen. Gemäß dem Hochschulgesetz NRW in der Fassung vom 1. Januar 2007 (HG NRW) § 7 Abs. 2 S. 1 sind die Hochschulen sogar zur Überprüfung der Qualität der Lehre verpflichtet: »Zur Qualitätsentwicklung und -sicherung überprüfen und bewerten die Hochschulen regelmäßig die Erfüllung ihrer Aufgaben, insbesondere im Bereich der Lehre.« Das HG NRW regelt in §7 Abs. 2 S. 4 auch die letztendliche Verwendung der Daten: »Die Ergebnisse der Evaluation sind zu veröffentlichen.« Der Kreis der Involvierten wird in Abs. 4 S. 1 bestimmt: »Alle Mitglieder und Angehörigen der Hochschule haben die Pflicht, an Akkreditierung und Evaluation mitzuwirken.«

Als Argument speziell gegen Onlinebefragungen wurde die höhere Beteiligungsquote bei Papierfragebögen, die in den jeweiligen Lehrveranstaltungen ausgefüllt werden, vorgebracht. Grundsätzlich ist hier auf die Problematik der Repräsentativität von Umfragen hinzuweisen, wie sie unter anderem bereits von Maurer & Jandura (2008), Pötschke (2008) sowie Riek, Krüger & Rudinger (2009) erörtert wurde. Tatsächlich lassen sich unter Verwendung von Papierfragebögen Beteiligungsquoten von bis zu 100% erzielen, sofern diese in den Lehrveranstaltungen ausgeteilt, sofort ausgefüllt und anschließend wieder eingesammelt werden. Dieser hohe Rücklauf kann unter gleichen Bedingungen auch bei Onlinebefragungen erreicht werden. Dazu müssen auch die Onlinebefragungen während der Anwesenheitszeit der Lehrveranstaltungen (beispielsweise in einem CIP-/PC-Pool) durchgeführt werden.

Dieses Vorgehen wird bereits erfolgreich an der Universität Bonn praktiziert. Ein deutlicher Nachteil der Paper-Pencil-Methode gegenüber dem Onlineverfahren ist die im Vergleich geringere Ökonomie. Insbesondere bei Vorlesungen mit mehreren hundert Teilnehmern ist der Auswertungsaufwand speziell bei Fragen mit offenem Antwortformat extrem groß und die Datenübertragung fehleranfällig. Ergebnisberichte können dann nur mit erheblichem personellen Aufwand und nach langer Auswertungszeit fertiggestellt werden. Entsprechend können qualitätsfördernde Maßnahmen erst sehr spät entwickelt und umgesetzt werden. Ein weiterer Vorteil von Onlinebefragungen liegt in der sicheren Filterführung bei adaptiven Fragebögen. Diese ist in der Papierform unkomfortabler zu handhaben und fehleranfälliger als bei einer Onlinebefragung. Des Weiteren findet beim Onlineverfahren zur Vermeidung von Eingabefehlern der Befragten eine Plausibilitätsprüfung in Echtzeit statt. Hinzu kommt ein spezifisches Problem bei den modularisierten Studiengängen: In verschiedenen Lehreinheiten der Universität Bonn können Lehrveranstaltungen gleichzeitig mehreren Modulen zugeordnet sein. Für die Modul- und Lehrveranstaltungsevaluation ist es jedoch wichtig zu wissen, innerhalb welchen konkreten Moduls ein Studierender eine Lehrveranstaltung bewertet. Dazu müssen Modul und Lehrveranstaltung eindeutig zugeordnet werden. In der Papierform kann diese Information nur durch die Studierenden selbst angegeben werden, sofern keine personalisierten Fragebögen ausgeteilt werden sollen. Die Ausgabe personalisierter Fragebögen wäre jedoch ein großes logistisches Problem und eine Anonymität der Befragten kaum vermittelbar bzw. umsetzbar. Insofern bleibt nur ein aufwendig auswertbares und bei der Eingabe fehleranfälliges Freitextfeld oder die logistisch aufwendige Variante, alle möglichen Module auf dem Papierfragebogen zur Auswahl anzubieten. Da es erfahrungsgemäß immer wieder vorkommt, dass Studierende nicht wissen, innerhalb welchen Moduls sie eine Lehrveranstaltung besuchen, hat das Onlinesystem des ZEM auch hier Vorteile. Dabei können nämlich die Studierenden eines Moduls über das elektronische Prüfungsverwaltungssystem ermittelt und per E-Mail zur Bewertung ihrer individuellen Module und Lehrveranstaltungen eingeladen werden. Die Studierenden sind gehalten, das dabei verwandte Universitätsemailkonto regelmäßig einzusehen.

Das vom ZEM entsprechend dem Bonner Modell der Hochschulevaluation eingesetzte webbasierte Onlinebefragungssystem weist also deutliche Vorteile gegenüber Befragungen mittels der Paper-Pencil-Methode auf.

Auch wenn die gegen die vom ZEM durchgeführte Modul- und Lehrveranstaltungsevaluation vorgebrachten Befürchtungen größtenteils entkräftet werden können, sollten die Bedenken aus den Lehreinheiten berücksichtigt und sorgfältig geprüft werden. Für die Entwicklung und Umsetzung qualitätsfördernder Maßnahmen ist die Akzeptanz des Verfahrens durch die Lehrenden und die Beteiligung aller Involvierten am Verfahren unbedingt erforderlich. Vorbehalte sollten in einem konstruktiven Dialog aufgelöst werden und

der potentielle Nutzen des Verfahrens herausgestellt werden: »Evaluation muß für die Beteiligten einen erkennbaren Nutzen bringen; andernfalls würde jeder von Ihr angestoßene Prozeß schnell wieder zum Stillstand kommen.« (Kromrey, 2006). Wie der Rektor der Universität Bonn, Prof. Dr. Jürgen Fohrmann, kürzlich in einem Zeitungsinterview mitteilte, soll an der Universität Bonn ein Belohnungssystem für gute Lehre etabliert werden (Seiler, 2009). Dieses setzt ein System für faire Vergleiche voraus, wie es die Modul- und Lehrveranstaltungsevaluation durch das ZEM bereits bietet.

Literatur

- DER Rektor der Rheinischen-Friedrich-Wilhelms-Universität Bonn (Hrsg.) (2008). *Neufassung der Evaluationsordnung für Lehre und Studium der Rheinischen Friedrich-Wilhelms-Universität Bonn (LEVO) vom 04. August 2008*. Bonn: Rheinische Friedrich-Wilhelms-Universität Bonn.
- HEIDEMANN, K., Rietz, C., Krahn, B., Riek, S. & Rudinger, G. (2005). Das Bonner Modell der Hochschulevaluation: Ein universitätsweites Verfahren zur Bewertung von Hochschulleistungen. *Zeitschrift für Evaluation*, 2, 195-219.
- KROMREY, H. (2006). Qualität und Evaluation im System Hochschule. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 234-259). Münster: Waxmann.
- LANDESGESETZ über die Hochschulen des Landes Nordrhein-Westfalen (Hochschulgesetz – HG) in der Fassung des Hochschulfreiheitsgesetzes vom 1. Januar 2007 (Hochschulfreiheitsgesetz – HFG).
- MAURER, M. & Jandura, O. (2008). Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung* (S. 61-73). Wiesbaden: VS.
- PÖTSCHKE, M. (2008). Potentiale von Online-Befragungen: Erfahrungen aus der Hochschulforschung. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung* (S. 75-89). Wiesbaden: VS.
- RIEK, S., Krüger, T. & Rudinger, G. (2009). Befragungsmethodische Aspekte. In K. Lichtblau & A. Neligan (Hrsg.), *Das IW-Zukunftspanel. Ziele, Methoden, Themen und Ergebnisse* (S. 35-50). Köln: Deutscher Instituts-Verlag.
- SEILER, J. (2009, 18. August). Jürgen Fohrmann: »Sanierung kostet 850 Millionen Euro«. *General Anzeiger Bonn*. Zugriff am 20. August 2009 unter <http://www.general-anzeiger-bonn.de/index.php?k=news&itemid=10028&detailid=628237>

Eva Sindern & Sandra Pietrangeli

Validitätsforschung im Zuge der Entwicklung fachspezifischer Online Self-Assessments an der Rheinischen Friedrich-Wilhelms-Universität Bonn

Hintergrund

Die Ansprüche der deutschen Hochschulen ebenso wie die Apelle aus Bildungs- und Wirtschaftspolitik sind eindeutig: Die Studienerfolgsquote soll deutlich gesteigert werden, gleichzeitig strebt man eine Senkung der durchschnittlichen Studiendauer und eine Niveausteigerung der Abschlussnoten an, um im internationalen Vergleich exzellent abzuschneiden. Dem gegenüber stehen anhaltend hohe Studienabbrecherzahlen. So bricht etwa ein Fünftel aller Studienanfänger das Studium ab, und durch die Einführung der Bachelor- und Masterstudiengänge hat sich die Abbruchquote sogar noch erhöht (vgl. Heublein, Schmelzer & Sommer, 2008).

Ursachen für Studienabbruch oder -fachwechsel finden sich, ebenso wie Gründe für mangelnden Studienerfolg und Studienunzufriedenheit, oftmals in einer mangelnden Passung zwischen den Individualprofilen von Studienanfängern (Kompetenzen, Interessen und Neigungen, Erwartungen, Ziele und Bedürfnisse) und dem (fachspezifischen) Anforderungsprofil der Hochschule (inhaltliche Anforderungen, strukturelle Anforderungen, Ziele und Verhaltenserwartungen) (vgl. Heublein, 2001; Lewin und Lischka, 2004).

Um dieser Problematik zu begegnen, ist eine Verbesserung der Informationsmöglichkeiten für Studieninteressierte über das Studium im Allgemeinen, die verschiedenen Studiengänge sowie Spezifika der jeweiligen Hochschule sinnvoll. Immer mehr Hochschulen im deutschsprachigen Raum begegnen dieser Aufgabe u.a. durch die Implementierung so genannter Online Self-Assessments (OSA). Eine einheitliche Definition für OSA gibt es nicht. Im Hochschulkontext lassen sie sich in Abgrenzung zu Verfahren der Studierendenauswahl als internetbasierte »Selbsttests zur Studienorientierung« (Hell, 2009) beschreiben, die Studieninteressierten die Möglichkeit bieten, die eigenen studienrelevanten Fähigkeiten und Kompetenzen zu erproben und über eine differenzierte Rückmeldung zu einer Einschätzung der eigenen Passung zu den Anforderungen für das Studium zu gelangen. Diese Einschätzung soll wiederum sog. *Selbstselektionsprozesse* anregen, d.h. Studieninteressierte mit einer hohen Passung sollen in ihrer Entscheidung für Studienfach und -ort bestärkt werden, während Aspiranten mit einer geringen Passung abgeschreckt werden sollen. Neben dem (eignungs-)diagnostischen

Aspekt können mittels Self-Assessment Informationen bezüglich der Studieninhalte und -ansprüche vermittelt werden. Die Studieninteressierten werden so im Sinne einer Realistic-Job-Preview (Wanous, 1989) über überhöhte oder falsche Vorstellungen und Erwartungen aufgeklärt und ein möglicher *Realitätsschock* kann vermieden bzw. abgemildert werden.

Umsetzung von OSA an der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gemäß der Ziel- und Leistungsvereinbarungen mit dem Ministerium für Innovation, Wissenschaft, Forschung und Technologie (MIWFT) des Landes Nordrhein-Westfalen hat die Universität Bonn zum Mai 2007 für erste Pilotfächer fachspezifische Online Self-Assessments für Bachelorstudiengänge entwickelt. Diese Pilotfächer sind Geodäsie; Germanistik, Vergleichende Literatur- und Kulturwissenschaft; Geschichte und Mathematik; in den nächsten Jahren werden weitere Fächer einbezogen. Eingebettet sind die Bonner OSA in ein neu konzipiertes Studierendeninformationssystem (SIS; online verfügbar unter www.studienscout.uni-bonn.de), das explizit auf die Bedürfnisse von Studieninteressierten ausgerichtet ist. Das SIS ist als zweistufiges Webportal (Eingangsseite) im Corporate Design der Universität Bonn gestaltet, eine zielgruppengerechte Ansprache mit hohem Wiedererkennungswert soll durch den virtuellen Studienberater *Academicus* erreicht werden (s. Abbildung 1). Auf der ersten Stufe des Webportals findet der Studieninteressierte allgemeine Informationen zum Universitätsleben, zur Stadt Bonn, zu Finanzierungsmöglichkeiten, Tipps für die Wohnungssuche etc. Auf der zweiten Stufe werden neben dem Zugang zum fachspezifischen Self-Assessment Tool auch weiterführende studiengangspezifische Informationen wie das Studiengangprofil, der -verlaufsplan, fachspezifische Beratungsangebote und Berufsperspektiven angeboten.

The screenshot shows the homepage of 'Studienscout Academicus' at the University of Bonn. At the top right, there are links for 'Übersicht', 'Barrierefreiheit', 'Kontakt', 'Impressum', and a search bar. The main navigation bar includes 'Die Universität', 'Einrichtungen', 'Studium', and 'Forschung'. A left sidebar lists various study-related categories like 'Studienangebot', 'Studienangebote A - Z', 'Grundständige Studiengänge', and 'Studienscout Academicus'. The main content area is titled 'Studienscout Academicus' and contains a welcome message, a cartoon character 'Academicus' pointing, and detailed text about the service's purpose: to inform students about study programs in Bonn and help them choose based on their interests and abilities. A right sidebar lists contact information for 'Ansprechpartner' and 'Academicus'.

Abbildung 1: Startseite des Studieninformationssystems »Studienscout Academicus« mit der Figur des virtuellen Studienberaters »Academicus«

Zur Programmierung der OSA wird die Open Source Software testMaker der RWTH Aachen verwendet (Milbradt, Zimmerhofer & Hornke, 2007). Voraussetzung für die Teilnahme an einem OSA ist die Registrierung mit einer gültigen Email Adresse. Nach Aktivierung des Testzugangs kann der Nutzer an allen veröffentlichten OSA teilnehmen, einen bereits begonnenen Selbsttest fortsetzen oder die Ergebnisse abgeschlossener Tests einsehen. Abbildung 2 zeigt die Eingangsseite zur Registrierung zu den Bonner Selbsttests.

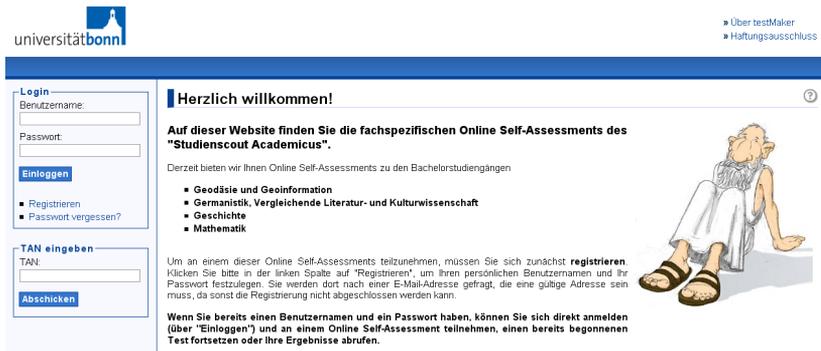


Abbildung 2: Allgemeine Startseite der Bonner OSA mit Registrierungsmöglichkeit

Die Bonner OSA beinhalten standardmäßig Leistungsaufgaben zu allgemeinen und fachspezifischen Studienkompetenzen sowie Selbsteinschätzungsfragen zu studienrelevanten Interessen und Erwartungen an Inhalte und Anforderungen des Studiums. Ergänzt werden kann dieses Angebot um Skalen zur Leistungsmotivation und Aufgaben zur Internetrecherche sowie um ein fachunspezifisches *Situational Judgment Inventory* (Weekley, & Ployhardt, 2006). Das Situational Judgment Inventory umfasst insgesamt neun *Typische Situationen aus dem Studienalltag* und eine variierende Anzahl an Handlungsalternativen.

Erste Ergebnisse zur Validität der OSA

Bei OSA handelt es sich ausdrücklich nicht um Verfahren zur Studierendenauswahl, sie haben vielmehr zum Ziel Studieninteressierte möglichst umfassend über Anforderungen und Herausforderungen eines (Fach-)Studiums zu informieren und über Selbstselektion die Grundquote geeigneter Studienbewerber zu erhöhen (Trost, 2005). Die Entwicklung von OSA als psychologisch-diagnostische Verfahren erfordert neben einer systematischen Anforderungsanalyse (vgl. Wissenschaftsrat, 2004; DIN, 2002) zur Ermittlung der relevanten Studienanforderungen auch eine Überprüfung der Verfahren an Studierenden und Schülern hinsichtlich der diagnostischen Gütekriterien, bevor die OSA im Internet veröffentlicht werden sollten (vgl. Hell, 2009; Kubinger, 2006a). Wichtige Kennziffern der Validität sind dabei der statistische Zusammenhang von kognitiven Leistungen im OSA mit der Abiturdurchschnittsnote und der Studiennote sowie von Ausprägungen der fachspezifischen Interessen und studienangabezogenen Erwartungen mit den *soften* Studienerfolgskriterien Studienzufriedenheit und Studienabbruch (vgl. Rindermann, 1999).

Zwischen März und April 2009 wurden die Bonner OSA an Bachelorstudierenden der Pilotfächer und an Oberstufenschülern getestet. Die Teilnahme

an der zweiwöchigen Testphase war freiwillig. Die Daten der Testpersonen wurden zur Bestimmung der kriterienbezogenen Validität herangezogen. Nicht für alle Testpersonen konnten eine Abiturdurchschnittsnote und/oder Studiennote(n) erhoben werden, so dass sich die Basis zur Berechnung der Validitätskoeffizienten teilweise erheblich reduzierte. Tabelle 1 gibt einen Überblick über die Anzahl der Teilnehmer in jedem Pilotfach insgesamt und getrennt nach Schülern und Studierenden, die das OSA *vollständig* und inkl. des Evaluationsfragebogens bearbeitet haben.

Tabelle 1: Teststichproben in den Pilotfächern (vollständige Bearbeitung des OSA inkl. Evaluationsfragebogen)

<i>Pilotfach</i>	<i>Gesamt (N)</i>	<i>Schüler (n₁)</i>	<i>Studierende (n₂)</i>
<i>Geodäsie und Geoinformation</i>	41	15	26
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft</i>	110	3	107
<i>Geschichte</i>	42	14	28
<i>Mathematik</i>	65	47	18

Anhand der Daten aus der Testphase sollten für jedes OSA folgende Fragen geklärt werden:

- (F1) Weisen die Verfahren eine *angemessene Schwierigkeit* auf?
- (F2) Hängt die Leistung im OSA statistisch mit der *Abiturdurchschnittsnote* zusammen?
- (F3) Besteht bei Studierenden ein *Zusammenhang zwischen Leistungen im OSA und Studienleistung*?
- (F4) Hängen studiengangbezogene *Interessen und Erwartungen* mit *Studienzufriedenheit* bzw. *Studienabbruch* zusammen?

Zur Beantwortung von (F1) wurden im Zuge einer Itemanalyse die Schwierigkeitsindizes der Aufgaben berechnet. Aufgaben, die i.S. des *weiten Kriteriums* nach Fisseni (2004) zu leicht oder zu schwer waren, wurden in Absprache mit den Fachvertretern bei der Überarbeitung zum Großteil aus den OSA entfernt. Einige besonders leichte Items wurden auf Grund ihrer Funktion als Eisbrecher (Fisseni, 2004) oder wegen ihres Informationsgehalts beibehalten. Die Schwierigkeiten der OSA für die Bearbeitung aller Leistungsaufgaben vor der Eliminierung von Items sind insgesamt und getrennt nach Schülern und Studierenden in Tabelle 2 dargestellt:

Tabelle 2: Gesamtschwierigkeit der OSA insgesamt und getrennt nach Schülern und Studierenden vor der Überarbeitung

<i>Pilotfach</i>		<i>n</i>	<i>Schwierigkeit (p)</i>	<i>SD</i>
<i>Geodäsie und Geoinformation</i>	Gesamt	41	.67	,102
	Schüler	15	.63	,102
	Studierende	26	.70	,094
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft¹</i>	Gesamt	110	.74	,102
	Schüler	3	—	
	Studierende	107	—	
<i>Geschichte</i>	Gesamt	42	.69	,100
	Schüler	14	.63	,111
	Studierende	28	.72	,080
<i>Mathematik</i>	Gesamt	65	.74	,162
	Schüler	47	.70	,164
	Studierende	18	.85	,089

¹Wegen der kleinen Stichprobe für Schüler wurden keine getrennten Schwierigkeitsindizes berechnet

Die Bestimmung des statistischen Zusammenhangs von der Leistung im OSA mit der Abiturdurchschnittsnote (F2) erfolgte, indem die Gesamtleistung¹ im OSA mit der gestuften Abiturdurchschnittsnote² korreliert wurde. Die Abiturdurchschnittsnote hat sich in zahlreichen Studien als zuverlässiger Prädiktor für Studienerfolg erwiesen und scheint somit als Kriterium geeignet (vgl. Rinderman, 1999). Für alle fachspezifischen OSA konnten positive Korrelationen zwischen der Gesamtleistung im OSA und der Abiturdurchschnittsnote gefunden werden, so dass die Verfahren im Hinblick auf dieses Kriterium als valide angesehen werden können (vgl. Tabelle 3).

1 Mittelwert aller kognitiven Aufgaben, die in »richtig/falsch« (1/0 codiert) klassifiziert werden können. Die Gesamtleistung kann Werte zwischen 0 und 1 annehmen und gibt folglich auch die Gesamtschwierigkeit des OSA an.

2 1=1,0 bis 1,4; 2=1,5 bis 1,9; 3=2,0 bis 2,4; 4=2,5 bis 2,9; 5=3,0 bis 3,4; 6=3,5 bis 4,0

Tabelle 3: Korrelation von Gesamtleistung im OSA mit der Abiturdurchschnittsnote

<i>Pilotfach</i>	<i>n</i>	<i>Spearman ρ</i>
<i>Geodäsie und Geoinformation</i>	28	.214
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft</i>	105	.261**
<i>Geschichte</i>	28	.411*
<i>Mathematik</i>	21	.659**

*Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Der Zusammenhang von Studienleistungen und der im OSA erbrachten Leistung wurde anhand der Korrelation von Gesamtleistung und bisherigen Studiennoten im Bachelorstudium bewertet. Während für die Studiengänge *Germanistik, Vergleichende Literatur- und Kulturwissenschaft* und *Geschichte* jeweils eine Studiennote erfragt wurde (Note im Studienportal bzw. Durchschnittsnote aus Studien- und Prüfungsleistungen aus Kern-, Begleitfach und Freiem Wahlpflichtbereich im Fachbereich Geschichte), wurden für die Studiengänge *Geodäsie und Geoinformation* und *Mathematik* drei relevante Studiennoten erfragt (Geodäsie und Geoinformation: Note₁ »Modul Ingenieurmathematik«, Note₂ »Modul Experimentalphysik«, Note₃ »Modul Einführung in Geoinformation und Kartographie«; Mathematik: Note₁ »Analysis 1«, Note₂ »Lineare Algebra 1«, Note₃ »Algorithmische Mathematik 1«). Für die Fächer *Geodäsie und Geoinformation* und *Mathematik* wurde zu Gunsten der Übersichtlichkeit aus den drei Einzelkorrelationen jeweils eine mittlere Korrelation berechnet (vgl. Bortz, 1999; S. 209-210). Es wird ersichtlich, dass auch für das Kriterium *Studiennote* zufriedenstellende Korrelationskoeffizienten ermittelt werden konnten, wie Tabelle 4 zeigt. Mit Ausnahme für das OSA *Germanistik, Vergleichende Literatur- und Kulturwissenschaft* zeigen sich ausgeprägte Zusammenhänge zwischen der Gesamtleistung im jeweiligen OSA und den Studienleistungen.

Tabelle 4: Mittlere Korrelation von Gesamtleistung im OSA mit Studiennoten

<i>Pilotfach</i>	<i>n</i>	<i>Spearman ρ</i>	
<i>Geodäsie und Geoinformation</i>	$n_1=22$ $n_2=22$ $n_3=11$	$\rho_1=.437^*$ $\rho_2=.239$ $\rho_3=.310$.336
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft</i>	101		.140
<i>Geschichte</i>	25		.499*
<i>Mathematik</i>	$n_1=16$ $n_2=16$ $n_3=14$	$\rho_1=.431$ $\rho_2=.287$ $\rho_3=.092$.285

*Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

Die Studienzufriedenheit (»Wie zufrieden sind Sie mit Ihrer Studienfachwahl?«) und beabsichtigter Studienabbruch (»Haben Sie bereits einmal darüber nachgedacht, Ihr Studienfach zu wechseln oder das Studium abzubrechen?«) wurden jeweils mit einer 4er-Skala³ erhoben. Der statistische Zusammenhang von studiengangbezogenen Interessen und Erwartung an Inhalte und Anforderungen des Studiums mit Studienzufriedenheit bzw. Studienabbruch (F4) wurde ebenfalls über die Berechnung von Korrelationskoeffizienten ermittelt (vgl. Tabelle 5a und 5b). Auffällig ist dabei, dass zwischen Erwartungen und späterer Studienzufriedenheit in den geisteswissenschaftlichen Fächern in der Teststichprobe kein bemerkenswerter statistischer Zusammenhang besteht, während sich ein signifikant positiver Zusammenhang von studiengangbezogenen Interessen und Studienzufriedenheit ergibt. Studienabbruch korreliert sowohl mit studiengangbezogenen Interessen als auch mit Erwartungen an das Studium moderat (Ausnahme *Germanistik, Vergleichende Literatur- und Kulturwissenschaft*).

3 Studienzufriedenheit: Sehr zufrieden – eher zufrieden – eher unzufrieden – sehr unzufrieden
Studienabbruch: Nein – ja, aber nicht ernsthaft – ja, ernsthaft – ja, ich werde das Fach wechseln bzw. das Studium abbrechen

Tabelle 5a: Korrelation von studiengangbezogenen Interessen mit Studienzufriedenheit und Abbruch

<i>Pilotfach</i>		<i>Studienzufriedenheit</i>	<i>Studienabbruch</i>
<i>Geodäsie und Geoinformation</i>	n	26	26
	Spearman ρ	.427*	.399*
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft</i>	n	107	107
	Spearman ρ	.362**	.212*
<i>Geschichte</i>	n	28	28
	Spearman ρ	.486**	.313
<i>Mathematik</i>	n	18	18
	Spearman ρ	.298	.308

*Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Tabelle 5b: Korrelation von *Erwartungen an Inhalte und Anforderungen des Studiums* mit Studienzufriedenheit und Abbruch

<i>Pilotfach</i>		<i>Studienzufriedenheit</i>	<i>Studienabbruch</i>
<i>Geodäsie und Geoinformation</i>	n	26	26
	Spearman ρ	.208	.452*
<i>Germanistik, Vergleichende Literatur- und Kulturwissenschaft</i>	n	107	107
	Spearman ρ	.078	.042
<i>Geschichte</i>	n	28	28
	Spearman ρ	.048	.247
<i>Mathematik</i>	n	18	18
	Spearman ρ	.260	.250

*Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

**Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Ausblick

Insgesamt können die Ergebnisse der Testphase als zufrieden stellend bezeichnet werden. Die Gesamtschwierigkeit der Testaufgaben war eher leicht, wobei die Aufgaben von Studierenden erwartungsgemäß häufiger gelöst werden konnten als von Schülern. Bei der Konstruktion von typischen Eignungstests wird eine Schwierigkeit von $p=0.5$ angestrebt, um geeignete von ungeeigneten Bewerbern möglichst perfekt trennen zu können (Fisseni, 2004). Da es sich bei OSA jedoch nicht um ein Auswahlverfahren handelt und die Nutzer nicht unnötig von der Schwierigkeit der Aufgaben frustriert und somit von einem Studium abgehalten werden sollen, ist eine Lösungswahrscheinlichkeit von über 50% hier durchaus akzeptabel. Vor der Veröffentlichung wurde die Mehrzahl der zu schwierigen oder zu leichten Aufgaben aus den OSA eliminiert, so dass die Schwierigkeit der veröffentlichten Versionen vermutlich höher ausfällt.

Die hohen Zusammenhänge im Hinblick auf das kapitale Kriterium *Abiturdurchschnittsnote* sprechen für die Validität der kognitiven Aufgaben. Die statistischen Zusammenhänge mit den Studienleistungen fallen mit Ausnahme für das Fach *Germanistik, Vergleichende Literatur- und Kulturwissenschaft* ebenso positiv aus. Auch die fachspezifischen Interessen stehen in einem ausgeprägten Zusammenhang mit der Studienzufriedenheit und dem Studienabbruch, so dass nicht nur die Leistungsindikatoren der jeweiligen OSA als zuverlässig angesehen werden können, sondern auch einer der nicht kognitiven Indikatoren der OSA mit den soften Kriterien für Studienerfolg korreliert. Die Betrachtung der Ergebnisse für den Zusammenhang von Er-

wartungen und soften Kriterien des Studienerfolgs fallen für die geisteswissenschaftlichen Fächer weniger zufriedenstellend aus. Zu beachten ist bei Interpretation der Ergebnisse zur Validität der Erwartungen, dass diese im Fall der Gruppe der Studierenden retrospektiv erfasst wurden, das heißt die Studierenden wurden gebeten, die Fragen zu *Erwartungen an das Studium* so zu beantworten, wie es ihrer Erwartungshaltung vor Beginn des Studiums entsprochen hätte. Insofern können hier Verzerrungs- und Gedächtniseffekte nicht ausgeschlossen werden. Um die Güte der Erwartungssitems umfassender zu prüfen, sind längsschnittliche Betrachtungen in der Stichprobe von *Echtnutzern* erforderlich.

Im Vergleich der fachspezifischen Verfahren untereinander fällt auf, dass für das OSA *Germanistik, Vergleichende Literatur- und Kulturwissenschaft* die niedrigsten Zusammenhänge ermittelt wurden. Das Fach weist an der Universität Bonn eine äußerst hohe Bandbreite an Inhalten auf, da im Zuge der Umstellung von Magister- auf Bachelorstudiengänge unterschiedliche Disziplinen in diesem Studiengang vereint wurden. So gestaltete es sich bereits bei der Anforderungsanalyse schwierig, allgemeingültige Anforderungen für Studienanfänger zu formulieren. Daher lag der Fokus bei der Konzeption der Aufgaben und Fragen in Absprache mit den Fachvertretern vermehrt auf der Darstellung der zahlreichen fachlichen Inhalte als in der Erfassung von insgesamt gültigen kognitiven Voraussetzungen für das Studium. Insofern wurde eine weniger gute Vorhersage von Leistungen zu Gunsten der Informationsfunktion in Kauf genommen. Dieser Umstand verdeutlicht, dass neben dem kriterienbezogenen Evaluationsansatz auch eine Überprüfung der Funktion der Informationsvermittlung durch OSA notwendig ist. Ergänzt werden sollte die Evaluation zusätzlich durch die Ermittlung der Akzeptanz in der Zielgruppe.

Aufgrund der beschriebenen Ergebnisse kann von der kriterienbezogenen Validität der untersuchten OSA ausgegangen werden, so dass diese nach einer Überarbeitung im Mai 2009 gemeinsam mit den begleitenden Informationen im SIS der Universität Bonn veröffentlicht wurden. Die Entwicklung der Verfahren ist auch nach der Onlinestellung nicht vollständig abgeschlossen, vielmehr sollen die Verfahren durch längsschnittliche Begleitforschung zu diagnostischer Güte und Bewertung in der Zielgruppe weiter verbessert und an die Bedürfnisse von Studieninteressierten angepasst werden.

Literatur

- BORTZ, J. (1999). *Statistik für Sozialwissenschaftler* (5. Auflage, S. 209-210). Berlin: Springer.
- DIN Deutsches Institut für Normung e.V. (2002). *Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen: DIN 33430*. Berlin: Beuth-Verlag.

- FISSENI, H.J. (2004). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- HELL, B. (2009). Selbsttests zur Studienorientierung: nützliche Vielfalt oder unnützer Wildwuchs? In G. Rudinger & K. Hörsch (2009), *Self-Assessment an Hochschulen: Von der Studienfachwahl zur Profilbildung* (Applied Research in Psychology and Evaluation Vol. 4, S. 9-19). Bonn: University Press. Göttingen: V&R unipress.
- HEUBLEIN, U. (2001). *Studienabbruch als Korrektur der Studienentscheidung – Gründe für den Studienabbruch*. Zugriff am 19.08.2009 unter <http://www.his.de/publikation/seminar/Tagung2001/Heublein.pdf>
- HEUBLEIN, U., Schmelzer, R. & Sommer, D. (2008). *Die Entwicklung der Studienabbruchquote an den deutschen Hochschulen. Ergebnisse einer Berechnung des Studienabbruchs auf Basis der Absolventenjahrgangs 2006* (HIS: Projektbericht Februar 2008). Hannover: HIS Hochschul-Informationen-System GmbH.
- KUBINGER, K.D. (2006a). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- LEWIN, D. & Lischka, I. (2004). *Passfähigkeit beim Hochschulzugang als Voraussetzung für Qualität und Effizienz von Hochschulbildung* (Arbeitsberichte 6'04). Wittenberg: Institut für Hochschulforschung an der Martin-Luther-Universität Halle-Wittenberg.
- MILBRADT, A., Zimmerhofer, A. & Hornke, L. F. (2007). *testMaker – a software for web-based assessments* [Computer software]. Aachen: RWTH Aachen University, Department of Industrial and Organizational Psychology.
- RINDERMANN, H. (1999). Auswahl von Studienanfängern durch Universitäten – Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für differentielle und diagnostische Psychologie*, 20 (3), 172-191.
- TROST, G. & Haase, K. (2005). Hochschulzulassung: Auswahlmodelle für die Zukunft. Eine Entscheidungshilfe für die Hochschulen. In Stifterverband für die deutsche Wissenschaft (Hrsg.), *Schriftenreihe der Landesstiftung Baden Württemberg* (6). Essen und Stuttgart: Landesstiftung Baden Württemberg.
- WANOUS, J. P. (1989). Installing a realistic job preview: Ten tough choices. *Personel Psychology*, 42 (1), 117-134.
- WEAKLEY, J.A. & Ployhart, R.E. (2006). An introduction to situational judgement testing. In J.A. Weekley & R.E. Polyhardt (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 1-10). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- WISSENSCHAFTSRAT (2004). *Empfehlungen zur Reform des Hochschulzugangs*. Zugriff am 19.08.2009 unter <http://www.wissenschaftsrat.de/texte/5920-04.pdf>

Maximilian Ponert & Katharina Hörsch

Evaluation von Change Management am Beispiel der Bonner Universitätsverwaltung

In sich immer schneller verändernden und dynamischen Umwelten einer globalisierten Welt wird von Organisationen eine kontinuierliche Anpassung verlangt, um langfristig erfolgreich und konkurrenzfähig zu bleiben. Das gilt sowohl für Profit-Organisationen wie Unternehmen, als auch zunehmend für Non-Profit-Organisationen wie Hochschulen. Bedingt durch das Hochschulrahmengesetz (HRG) und die Hochschulgesetze der Länder müssen Hochschulen sich auch im Hinblick auf ihre Effizienz, Wettbewerbsfähigkeit und Sicherstellung von Qualität in Lehre und Forschung wandeln. Die Universität Bonn befindet sich, wie viele andere bundesdeutsche Universitäten, infolge von Gesetzen und Zielvereinbarungen derzeit und in den kommenden Jahren in einem kontinuierlichen Veränderungs- und Restrukturierungsprozess.

In Nordrhein-Westfalen wurden aus den bisher dem Land unterstellten Universitäten infolge des Inkrafttretens des Hochschulfreiheitsgesetzes (HFG) selbstverwaltete Körperschaften des öffentlichen Rechts. Die neugewonnene Autonomie im Bezug auf Finanz-, Personal- und Organisationsentscheidungen brachte Universitäten in die Situation, von nun an »unternehmerisch« agieren zu müssen. Darüber hinaus führte das Gesetz zur Sicherung der Finanzierungsgerechtigkeit im Hochschulwesen (HFGG) zu neuen Verwaltungsaufgaben bezüglich Verwaltung und Verwendung der erhobenen Studienbeiträge. Eine weitere Herausforderung für die deutschen Universitäten war die Umstrukturierung der Studiengänge auf Bachelor und Master gemäß dem 1999 in Kraft getretenen Bologna-Prozess.

Die Exzellenzinitiative von Bund und Ländern schließlich stellte die Universität Bonn in direkten Wettbewerb mit allen deutschen Hochschulen um die insgesamt ausgelobten 1,9 Mrd. Euro zur Förderung universitärer Spitzenforschung. Neben der Förderung ausgewählter Graduiertenschulen und Exzellenzcluster wurden die Zukunftskonzepte einzelner Universitäten mit jeweils 21 Mio. Euro pro Jahr gefördert. Das Zukunftskonzept wurde definiert als: »[...] ein Gesamtkonzept, das die langfristige strategische Ausrichtung der Spitzenforschung an der Universität plausibel macht, bestehend aus einer Profilbeschreibung für Forschung und Nachwuchsförderung sowie einer die eigenen Stärken und Schwächen berücksichtigenden Entwicklungsplanung.« (Deutsche Forschungsgemeinschaft & Wissenschaftsrat, 2005, S. 3). Auch wenn die Universität Bonn im Hinblick auf das Zukunftskonzept in

beiden Förderrunden 2007 nicht berücksichtigt wurde, positionierte sie sich im Zuge des Exzellenzwettbewerbs als »international operierende, kooperations- und schwerpunktorientierte Forschungsuniversität« (Universität Bonn, 2008) und entwickelte mit dem Hochschulentwicklungsplan ein universitäres Gesamtkonzept. Auf der Ebene der Universitätsverwaltung mündete dieses Konzept in den Verwaltungsentwicklungsplan mit dem Ziel, optimale institutionelle Bedingungen für Forschung und Lehre zu schaffen und bestehende und zukünftige Exzellenzinitiativen der Universität Bonn zu unterstützen. Die dazu erforderlichen Maßnahmen wurden in der Verwaltungsstrategie in insgesamt 30 Projekte festgeschrieben.

In der Universitätsverwaltung wurden Projekte bis zu diesem Zeitpunkt mit Hilfe eines intern verfassten Leitfadens bearbeitet, darüber hinaus wurden die Projektbeteiligten von der Stabstelle Projektmanagement in ihrer Arbeit unterstützt. Für die Vielzahl an neuen, teils dezentral organisierten und komplexen Projekten der Verwaltungsstrategie war dieses Vorgehen nicht mehr ausreichend. Die Verwaltung benötigte ein Instrumentarium, das es ermöglichte, die zu bearbeitenden Projekte und resultierenden Maßnahmen nach einheitlichen Standards abzubilden und, angepasst an das Informationsbedürfnis der jeweiligen Hierarchieebene, zu steuern und zu überwachen. Zusätzlich wurde es als erforderlich angesehen, die Herangehensweise an Projektarbeit und -management grundlegend zu verändern und zu professionalisieren.

Im Rahmen eines Einführungsprojekts wurde zunächst ein Best-Practice Projekt in jedem der acht Dezernate nach der neuen Methode bearbeitet. Die dafür erforderlichen Mitarbeiter arbeiteten eng mit der Stabstelle Projektmanagement, einer eingerichteten Projektgruppe sowie einem Lenkungsausschuss zusammen und fungierten gleichzeitig als Multiplikatoren, die ihren Kolleginnen und Kollegen die neue Herangehensweise näher bringen sollten. Zusätzlich traten die Dezernenten als Promotoren des Veränderungsprozesses auf, indem sie aktiv den Veränderungsprozess mit trugen. Im Projektverlauf wurden alle mit Projektarbeit betrauten Verwaltungsmitarbeiter im Umgang mit der neuen Methode geschult, mit dem Ziel, zum Projektabschluss die Bearbeitung aller Projekte der Verwaltungsstrategie mit dem Projektmanagementsystem zu erreichen.

Das Zentrum für Evaluation und Methoden (ZEM) evaluierte dieses Einführungsprojekt begleitend. Die zentrale, durch die Evaluation zu klärende Frage war, ob die Einführung des neuen Systems als erfolgreich einzustufen sei. Um den Veränderungsprozess mit seinen Auswirkungen und den Erfolg des neuen Systems beurteilen zu können, wurden die Verwaltungsmitarbeiter zu mehreren Zeitpunkten befragt.

Eine Organisation von einem bisherigen in einen neuen gewünschten Zielzustand zu überführen, ist eine Anstrengung, die in vielen Fällen scheitert oder nicht zum gewünschten Resultat führt (z.B. IBM, 2007). Wie Kieser

(1998) feststellt, finden oftmals keine wirklichen Veränderungen statt oder das Verhalten der Organisationsmitglieder ändert sich nicht nachhaltig. Oftmals spielen auch mikropolitische Interessen eine Rolle bei der abschließenden Bewertung des Veränderungsprozesses als Erfolg oder Misserfolg. Doyle, Clayton & Buchanan (2000) berichten, dass Veränderungsprozesse nur selten von einer systematischen Evaluation begleitet werden. Der Umgang mit Veränderungen in Unternehmen, aber auch in Hochschulen, beinhaltet neben dem Bewusstsein für Veränderung und dem zielgerichteten Einsatz von Fähigkeiten auch die Ableitung konkreter Maßnahmen (IBM, 2008). Damit die richtigen Maßnahmen bereits während des Veränderungsprozesses eingeleitet und zur Steuerung genutzt werden konnten, wurde die Einführung des Projektmanagementsystems evaluierend begleitet. Das Ziel der begleitenden Evaluation des Einführungsprozesses des elektronischen Projektmanagementsystems war u.a. auch, den Mitarbeitern eine Möglichkeit zur Meinungsäußerung und Partizipation zu geben. Durch die Bewertung des Einführungsprozesses zu mehreren Befragungszeitpunkten sollten Stärken und Schwächen im Verlauf des Projekts identifiziert und Maßnahmen zur Optimierung abgeleitet werden. In Zusammenarbeit mit Vertretern aus der Universitätsverwaltung wurden auf Basis des übergeordneten Evaluationsziels Themen für die Befragung formuliert. Im Vordergrund standen dabei die Wahrnehmung der Projektarbeit vor der Umstellung, die Einstellung der Mitarbeiter in Bezug auf den Veränderungsprozess, Ängste und Erwartungen sowie die Wahrnehmung und Bewertung der Projektarbeit mit dem neuen System.

Der Einführungsprozess wird als erfolgreich eingestuft, wenn Planung und Umsetzung von Projekten und Maßnahmen strukturiert ablaufen, der gegenseitige Informationsaustausch in der Zusammenarbeit der Projektmitarbeiter reibungsloser gestaltet wird und die eingeleiteten Maßnahmen zur Steuerung von Projekten nachhaltiger wirken.

Die Umsetzung der onlinegestützten Befragung erfolgte über eine professionelle Software für Umfrageforschung (SPSS Dimensions), die auf gesicherten Servern des ZEM verwaltet wird. Diese Software erlaubt eine automatisierte Befragungsdurchführung im Sinne eines standardisierten Befragungsmanagements. Einladungen zur Befragung und Reminder werden automatisiert per E-Mail verschickt. Dabei werden die Kontaktdaten der Verwaltungsmitarbeiter getrennt von den Antworten der Befragten in separaten Datenbanken erfasst. Eine Verknüpfung der Kontaktierungsdaten mit den Antworten der Befragten ist dadurch ausgeschlossen. Zur Regulierung des Teilnehmerkreises erhielten die Verwaltungsmitarbeiter individuelle Passwörter, die verhindern, dass die Befragung von nicht autorisierten Personen ausgefüllt wird und eine einmalige Teilnahme an der Befragung ermöglichen. Nach Abschluss der Befragung werden die Kontaktdaten automatisch gelöscht.

Neben der höheren Standardisierung und der Einfachheit der Durchführung hat die onlinegestützte Befragung weitere Vorteile gegenüber der Paper-Pencil- Variante: Die einfache Verwendung von Filterführungen ermöglicht einen individualisierten Befragungsverlauf (Porst, 1998). Online-Umfragen werden als »am wenigsten aufdringlich wahrgenommen« (Mühlenfeld, 2004, S. 60), so dass Antworten weniger Tendenzen sozialer Erwünschtheit aufweisen (Mühlenfeld, 2004).

Ferner werden die Daten automatisch in eine Auswertungsdatei übertragen, so dass Fehler durch die manuelle Eingabe, z.B. ungültige Werte, weitestgehend minimiert werden. Die Annahme, dass die Anzahl der falschen Angaben in onlinegestützten Fragebögen höher ist, hat sich nicht bestätigt (Bortz & Döring, 1999).

Inhalte der Befragungen zu den verschiedenen Zeitpunkten waren u.a. Informationsaustausch, Transparenz, Arbeitsabläufe, Arbeitskapazitäten, Ziele der Arbeit, Einhaltung von Terminen und Vereinbarungen, Umsetzung von Beschlüssen und Maßnahmen, Einschätzung der Veränderung verschiedener Aspekte im Arbeitsalltag, Erwartungen, Ängste, Gesamtbewertung der Universität Bonn, Bedienbarkeit der Software, Hilfestellungen und Auswirkung des Systems auf die eigene Arbeit. Neben quantitativen Fragen gab es auch Möglichkeiten zur freien Eingabe.

412 Verwaltungsmitarbeiter wurden zu den vier Befragungszeitpunkten per E-Mail eingeladen, an der Befragung teilzunehmen. Im Vorfeld der Befragungen wurden alle Mitarbeiter per Rundschreiben von der Universitätsleitung über das Evaluationsvorhaben informiert und zur Teilnahme aufgefordert.

Die erste Befragung (T-Null) wurde im August 2007 durchgeführt, zu diesem Zeitpunkt hatte das Einführungsprojekt gerade begonnen. Zur Steigerung der Rücklaufquote wurden im Befragungszeitraum zwei Reminder per E-Mail an bisherige Non-Responder versandt. 211 Personen haben diese erste Befragung vollständig abgeschlossen, das entspricht einem Rücklauf von 52,49%.

Die zweite Befragung (T-Eins) fand im November und Dezember 2007 statt. Zur Steigerung der Rücklaufquote wurde ein Reminder per E-Mail an alle Befragungsteilnehmer versandt. Auf die Versendung ausschließlich an Non-Responder wurde aus Gründen der Akzeptanzsteigerung verzichtet. 118 Personen haben die zweite Befragung vollständig abgeschlossen, das entspricht einem Rücklauf von 29,35%.

Die dritte Befragung (T-Zwei) fand im März 2008 kurz nach Abschluss des Einführungsprojektes statt. Auf Reminder wurde aufgrund des kurzen Befragungszeitraumes von einer Woche verzichtet. 129 Personen haben die Befragung vollständig abgeschlossen, das entspricht einem Rücklauf von 32,08%.

Die Follow-up Befragung (T-Drei) wurde im November 2008 durchgeführt. Während des Befragungszeitraums wurden zwei Reminder verschickt. Insgesamt haben 174 Personen die Befragung abgeschlossen, das entspricht einem Rücklauf von 39,5%. Einen Überblick über die Befragungszeitpunkte gibt Tabelle 1.

Tabelle 1: Die vier Befragungszeitpunkte im Überblick

	<i>T-Null</i>	<i>T-Eins</i>	<i>T-Zwei</i>	<i>T-Drei</i>
<i>Größe der Stichprobe</i>	N=412			
<i>Befragungszeitraum</i>	02.08.- 20.08.2007	16.11.- 21.12.2007	17.03.- 21.03.2008	03.11.- 25.11.2008
<i>Reminder</i>	09.08.2007 und 16.08.2007 an Non- Responder	04.12.2007 an alle Be- fragten		13.11.2008 und 22.11.2008 an alle Befragten
<i>Rücklauf</i>	n ₀ =211 (52,49%)	n ₁ =118 (29,35%)	n ₂ =129 (32,08%)	n ₃ =174 (39,5%)

Die starke Sensibilisierung bei Verwaltungsmitarbeitern im Hinblick auf die Wahrung ihrer Anonymität hat dazu geführt, dass nach der ersten Befragung (T-Null) die Strategie des Reminderversands geändert wurde. Ursprünglich war vorgesehen, dass nur Mitarbeiter, die sich noch nicht an der Befragung beteiligt hatten, eine Erinnerungs-E-Mail erhalten. Diese Möglichkeit besteht, da die verwendete Umfragesoftware eine Loginkontrolle erlaubt, sodass bereits Befragte nicht erneut angeschrieben werden müssen. Da diese Prozedur automatisch abläuft, wird die Anonymität der Befragten zu keinem Zeitpunkt verletzt. Dennoch erhielten ab der zweiten Befragung alle Mitarbeiter Erinnerungs-E-Mails unabhängig davon, ob sie bereits an der Befragung teilgenommen hatten. Um trotz des automatisierten Versands der Passwörter an die Mitarbeiter keinen Zweifel an der Anonymität aufkommen zu lassen, wurde eine Beschreibung des Vorgehens im Intranet publiziert. Bei Fragen zu Anonymität und Datenschutz hatten die Befragten die Möglichkeit, zusätzliche Informationen zum Vorgehen zu erhalten.

Aufgrund der nicht auszuräumenden Angst vor Identifikation der eigenen Angaben im Fragebogen, wurde ab der zweiten Befragung zudem auf die automatisierte individuelle Anrede der Befragten verzichtet.

Die Motivation der Mitarbeiter zur Teilnahme an der Befragung gestaltete sich nicht leicht, da die Anzahl der Mitarbeiter, die tatsächlich mit dem System arbeiteten, zu Beginn nur langsam gesteigert wurde und viele Mitarbeiter keine Berührungspunkte mit dem Projektmanagementsystem hatten. Die Teilnahme an Befragungen hängt vornehmlich davon ab, wie wichtig die Befragungsinhalte für die Befragten sind (Pötschke, 2009), und ob sich aus den Befragungsergebnissen unmittelbare Konsequenzen ergeben. Trotz eines Informationsschreibens der Universitätsleitung, Gesprächen mit Dezenten und Erinnerungs-E-Mails haben sich vermutlich überwiegend die Mitarbeiter beteiligt, die bereits mit dem neuen System arbeiteten oder darin geschult waren.

Die Evaluation hat über alle vier Befragungszeitpunkte unterschiedliche Ergebnisse zutage gefördert. Zu jedem der vier Befragungszeitpunkte wurde auch die Bereitschaft der Weiterempfehlung sowohl der Universität Bonn als Arbeitgeber als auch als Hochschule erfasst. Ein Überblick dieser Ergebnisse findet sich im Beitrag von Hörsch & Rudinger zum Net Promoter Score® in diesem Band.

Im Vergleich der vier Befragungen wird deutlich, dass weiterhin Handlungsbedarf besteht, um die Einführung und Anwendung des Projektmanagementsystems nachhaltig zu festigen, Optimierungsbedarf zu erkennen und die Bindung der Mitarbeiter an die Universitätsverwaltung zu stärken.

Erkennbare Verbesserungen sind in den Bereichen Besprechungen, Transparenz von Abläufen und Verantwortlichkeiten zu verzeichnen. Hier zeigt sich bereits der positive Einfluss von strukturierter und standardisierter Abbildung von Projekten mit ihren Zielen, Zuständigkeiten und notwendigen Materialien im Projektmanagementsystem. Der Nutzen des Systems ist in diesen Bereichen für die Mitarbeiter bereits erkennbar. Insbesondere haben Besprechungen einen höheren Stellenwert, sie werden häufiger genutzt und als Informationsquelle erkannt. Ebenso sind Kollegen weiterhin erste Ansprechpartner, wenn es um Hilfestellungen für die Anwendung geht. Diese Art der zweiseitigen Kommunikation und Hilfestellung gilt es mindestens beizubehalten und wenn möglich zu erweitern.

Die Softwarelösung und ihre Bedienbarkeit stoßen auf Kritik, da die Software nach Meinung einiger Mitarbeiter für den Einsatz in einer Universitätsverwaltung nur eingeschränkt brauchbar ist. Hier besteht noch Optimierungsbedarf, ebenso im routinierten Umgang mit dem Projektmanagementsystem, durch den sich zum Beispiel die Einschätzung von Kapazitäten der beteiligten Mitarbeiter langfristig realistischer vornehmen lässt.

Bisher wird das Projektmanagementsystem für die eigene Arbeit tendenziell als mittelmäßig bewertet, was vermutlich im wahrgenommenen zeitlichen Mehraufwand im Zuge des Veränderungsprozesses und in den Schwächen der Softwarelösung begründet liegt. Durch zunehmende Routine im Umgang mit dem System werden möglicherweise mehr Mitarbeiter positive

Auswirkungen durch das System auf die eigene Arbeit erkennen. Nicht alle Effekte der Nutzung des Projektmanagementsystems lassen sich zum Ende der Einführungsphase erkennen. Erst durch eine längerfristige Nutzung des Systems wird sich abschließend zeigen, ob positive Veränderungen auch nachhaltig wirken und weitere Verbesserungen eingetreten sind. Zur Erkennung von Auswirkungen dieser langfristigen und routinierten Systemnutzung ist eine weitere, zeitverzögerte Follow-Up Befragung notwendig.

Auch aufgrund der Evaluationsergebnisse kann der Einführungsprozess grundsätzlich als erfolgreich eingestuft werden, da Planung und Umsetzung von Projekten und Maßnahmen strukturierter ablaufen, der gegenseitige Informationsaustausch in der Zusammenarbeit der Projektmitarbeiter reibungsloser gestaltet wird und die eingeleiteten Maßnahmen zur Steuerung von Projekten nachhaltiger wirken.

Letztlich sollten die Evaluationsergebnisse aber mit anderen Daten wie der Anzahl der Projekte und der Anzahl der Mitarbeiter, die das Projektmanagementsystem nutzen, in Bezug gesetzt werden, um den Erfolg der Einführung des Systems letztlich aus Sicht der Verantwortlichen zu bewerten.

Aus Sicht der Evaluatoren ist das Evaluationsziel erreicht, jedoch zeigt sich, dass organisationaler Wandel in einer Universität ein langwieriger Prozess ist. Die erfolgreiche Umsetzung von Veränderungen ist offenbar nur erfolgreich durchführbar, wenn Mitarbeiter ausreichend informiert sind und die Veränderung unterstützen. Zugleich ist es wichtig, die Belange und Bedürfnisse der Mitarbeiter nicht aus den Augen zu verlieren und z.B. durch eine begleitende Evaluation Hinweise auf Maßnahmen zur Verbesserung einzufordern. Transparenz von Beginn an und im Hinblick auf die Zielsetzung fördert nicht nur die inneruniversitäre Kommunikation, sondern stärkt damit auch die Loyalität der Mitarbeiter (Schüller, 2008).

Literatur

- BORTZ, J. & Döring, N. (1999). *Forschungsmethoden und Evaluation* (3. Auflage). Berlin: Springer.
- DEUTSCHE Forschungsgemeinschaft & Wissenschaftsrat. (2005). *Leitfaden für Zukunftskonzepte zum projektbezogenen Ausbau universitärer Spitzenforschung*. Zugriff am 04.09.2009 unter http://www.wissenschaftsrat.de/texte/exini_leitfaden.pdf
- DOYLE, M., Clayton, T. & Buchanan, D. (2000). Mixed results, lousy process the management experience of organisational change. *British Journal of Management*, 11, 59-80.
- IBM. (2007). *Making Change Work. Erfolgsfaktoren für die Einführung von Innovationen*. Zugriff am 08.09.2009 unter http://www-05.ibm.com/de/pressroom/downloads/mcw_2007.pdf

- IBM. (2008). *Making Change Work. Fortsetzung des Dialogs über das Unternehmen der Zukunft*. Zugriff am 08.09.2009 unter http://www-935.ibm.com/services/de/bcs/pdf/2009/making_change_work.pdf
- KIESER, A. (1998). Über die allmähliche Verfertigung der Organisation beim Reden. Organisieren als Kommunizieren. *Industrielle Beziehungen*, 5, 45-75.
- MÜHLENFELD, H.-U. (2004). *Der Mensch in der Online-Kommunikation*. Wiesbaden: Deutscher Universitäts-Verlag.
- PÖTSCHKE, M. (2009). Potentiale von Online-Befragungen: Erfahrungen aus der Hochschulforschung. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung*. S. 75-89 Wiesbaden: VS Verlag für Sozialwissenschaften.
- RHEINISCHE Friedrich-Wilhelms-Universität Bonn. (2008). *Das Leitbild*. Zugriff am 08.09.2009 unter http://www3.uni-bonn.de/die-universitaet/leitbild/unibonn/die-universitaet/leitbild/UBo_Leitbild_09_5-12-08.pdf
- SCHÜLLER, A.M. (2008). *Kundennähe in der Chefetage. Wie Sie Mitarbeiter kundenfokussiert führen*. Zürich: Orell Fuessli.

Katharina Hörsch & Georg Rudinger

Der Net Promoter Score®¹ an der Universität Bonn

Seit in den neunziger Jahren erstmals das Konzept des »Net Promoter Score« aufkam, hat dieser neue und erstaunlich einfache Indikator für das so schwer greifbare Konstrukt der »Kundenloyalität« zu vielen und zum Teil heftigen Diskussionen unter Loyalitätsforschern geführt. Was hat es mit der einfachen Frage nach der Wahrscheinlichkeit der Weiterempfehlung einer Institution oder eines Produkts an Freunde, Verwandte und Bekannte auf sich, die allen altgedienten, ausgeklügelten Marktforschungsstrategien überlegen sein soll?

Auf einer Skala von 0 bis 10, wie wahrscheinlich ist es, dass Sie XY an einen Freund, Verwandten oder Bekannten weiterempfehlen? ist die Frage der Wahl, mit der sich der Experte für Loyalitätsforschung, Fred Reichheld, einen Namen gemacht hat (Reichheld, 2003). Personen, die diese Frage mit Werten von 0 bis 6 beantworten, werden als *Kritiker (detractors)* bezeichnet. Von ihnen ist keine Weiterempfehlung zu erwarten, vielmehr zeichnet sich diese Gruppe dadurch aus, dass sie in ihrem Umfeld negative Mundpropaganda betreibt und die Institution oder das Produkt somit in Misskredit bringt. Eine Antwort von 7 oder 8 klassifiziert die entsprechenden Personen als *passiv Zufriedene (passively satisfied/passives)*: Weder besondere Ablehnung, noch besondere Zuneigung wird der Institution oder dem Produkt entgegengebracht. *Promotoren (promoters)* schließlich beantworten die Frage nach der Wahrscheinlichkeit einer Weiterempfehlung mit 9 oder 10 und werden somit als starke Befürworter des Produkts oder der Institution betrachtet (Reichheld, 2003). Die Einteilung in diese drei Kundensegmente verdeutlicht Abbildung 1.

¹ Net Promoter, NPS und Net Promoter Score sind registrierte Marken von Satmetrix Systems, Inc., Bain & Company und Fred Reichheld.

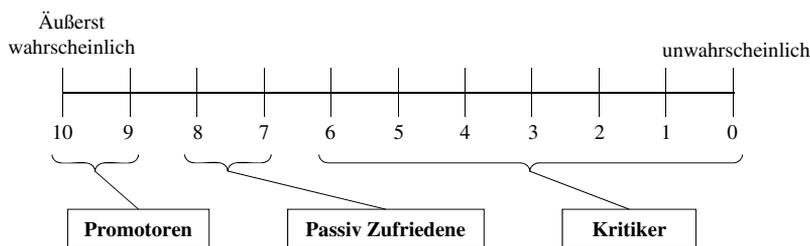


Abbildung 1: Einteilung in die drei Kundensegmente in Abhängigkeit vom Antwortverhalten (Reichheld & Seidensticker, 2006, S. 29)

Reichheld (2003, 2006) geht davon aus, dass Promotoren und Kritiker antagonistisch aufeinander wirken, d.h. der wirtschaftliche Nutzen, der einer Institution aus dem werbenden Verhalten der Promotoren entsteht, wird durch die negative und rufschädigende Haltung der Kritiker dezimiert. Der »wahre Wert« der Promotoren kann folglich immer nur unter der Berücksichtigung der Kritikerrate bestimmt werden (Reichheld & Seidensticker, 2006, S. 42). Dieser Logik folgend, berechnet sich die Kennziffer Net Promoter Score als Differenz aus dem prozentualen Anteil der Promotoren – P und dem der Kritiker – K: $(P - K = NPS)$. Der Net Promoter Score ist ein Maß für die *Qualität der Kundenbeziehung*. Reichheld & Seidensticker (2006), aber auch Keiningham, T. L., Cooil, B., Andreassen, T. W. & Aksoy, L. (2007) konnten nachweisen, dass Unternehmens- und Branchenwachstum in engem Zusammenhang mit einem hohen Net Promoter Score stehen.

»Wenn eine Organisation sich die Loyalität ihrer Mitglieder verdient, vergrößert sie die Ressourcen, die ihnen zur Verfügung stehen, und kann mehr Einfluss auf die Gesellschaft nehmen. Durch Nachverfolgung der Net Promoter Scores können die Anführer und Mitglieder der Gemeinschaft besser erkennen, wer die beste Arbeit leistet und um wen man sich kümmern muss. Die strikte Messung, Bekanntgabe und Handhabung einer einzigen Zahl – Prozentsatz der Promotoren unter den Mitgliedern minus Prozentsatz der Kritiker – ermöglicht es Organisationen, Best-Practices zu ermitteln, frische Ideen freizusetzen, Engagement aufzubauen und Verantwortlichkeit für die Ergebnisse zu etablieren. Alle Organisationen können wachsen und produktiver werden, indem sie die Zahl der Promotoren steigern und die Zahl der Kritiker senken.« (Reichheld & Seidensticker, 2006, S. 177)

Ob der Net Promoter Score sich aber auch im Non-Profit Bereich bewährt, und inwiefern er in anderen Unternehmenskulturen angewendet werden kann, soll in den folgenden Überlegungen mit Bezug auf erste Erfahrungen an der Universität Bonn erörtert werden.

Studenten oder Weiterbildungsteilnehmer als »Käufer« verschiedener Bildungsleistungen zu betrachten, bzw. treffender als »Klienten«, die das ihrige zum Erfolg ja beitragen müssen, stellt einen Paradigmenwechsel von hoch-

schulstrategischer Bedeutung dar. Insbesondere durch den zunehmenden Wettbewerb zwischen den Hochschulen wird das Bewusstsein für Studierende und Absolventen als Kunden (Schwaiger, 2003) der Universität vergrößert. Hochschulen verstehen sich immer mehr als »Dienstleister in Sachen Lehre bzw. Forschung« (Langer, Ziegele & Hennig-Thurau, 2001). Mit dieser veränderten Sichtweise geht die Notwendigkeit einher, diese Klienten an Hochschulen oder andere Bildungsträger zu binden und somit einen langfristigen Beitrag zur Profilbildung und Ausrichtung der Bildungseinrichtung zu leisten. Bindungsmanagement stellt damit auch an der Universität zunehmend eine wichtige Aufgabe dar.

Aber auch die Loyalität der Mitarbeiter wird mit wachsendem Wettbewerb im Hochschulbereich als Erfolgspotenzial erkannt. Nach Reichheld & Seidensticker (2006, S. 146) ist es nicht möglich, »die Loyalität der Kunden zu erringen, ohne sich vorher die Loyalität der eigenen Beschäftigten an der Kundenschnittstelle zu verdienen.«

Gute Beziehungen zwischen Unternehmen und ihren Kunden bringen Unternehmen wirtschaftliche Vorteile (Reichheld & Seidensticker, 2006). Analog dazu ist anzunehmen, dass gute Beziehungen zwischen der Universität und ihren Mitarbeitern, aber auch zwischen der Universität und ihren Studierenden und ehemaligen Studierenden, Vorteile in der Konkurrenz innerhalb der Hochschullandschaft bringen. Die daraus entstehenden Gemeinschaften oder Netzwerke, wie z.B. Alumni-Vereinigungen, bringen beiden Seiten Vorteile (Reichheld & Seidensticker, 2006).

Durch die Messung der Beziehungsqualität mit dem Net Promoter Score bei Studierenden, Absolventinnen und Absolventen sowie Mitarbeiterinnen und Mitarbeitern wird die Identifikation der »richtigen« Interaktionen und Investitionen für die langfristige Bindung möglich.

Seit 2007 wird die Frage nach der Weiterempfehlungsbereitschaft in verschiedenen Befragungen an der Universität Bonn als Indikator für die Bindung des einzelnen Studierenden an die Universität gestellt. Dieses Vorgehen wird beispielsweise durch die Studie »Hochschulbindung durch Student Services« von Lenecke (2005) gestützt, in der Hochschulbindung mittels Fragen zur Weiterempfehlungsbereitschaft des eigenen Studiengangs bzw. der Hochschule, Kontaktbereitschaft mit dem Fachbereich bzw. Eintritt in eine Alumnivereinigung und Bereitschaft, die Hochschule bzw. den Studiengang an dieser Hochschule wieder zu wählen, gemessen wurde. Zudem haben auch Langer, Ziegele & Hennig-Thurau (2001) Hochschulbindung u.a. als Weiterempfehlungsbereitschaft operationalisiert.

Ziel der Befragungen an der Universität Bonn ist die Erfassung der Wahrnehmung der Universität aus verschiedenen Perspektiven. Dafür wurden Studierende, Absolventinnen und Absolventen sowie Verwaltungsmitarbeiterinnen und -mitarbeiter befragt, wie wahrscheinlich es ist, dass sie die Universität Bonn weiterempfehlen. Denn nur die Betroffenen selbst können

Aufschluss über ihre Zufriedenheit als Studierende, Absolventen und Wissenschaftler mit ihrer Alma mater geben: Die Konkurrenz um gute Köpfe wächst, und zufriedene »Kunden« sind die beste Werbung.

Allgemeine Studierendenbefragung

Für die allgemeine Studierendenbefragung werden alle Studierenden an der Universität jährlich zum Ende des Wintersemesters eingeladen, Studium und Lehre an der Universität zu bewerten. Inhalte dieser Befragung sind neben Rahmenbedingungen des Studiums Engagement, Ausstattung und studienbegleitenden Aktivitäten auch das Image der Universität und Gründe für das Studium an der Universität Bonn.

Zum Wintersemester 2006/2007 wurde die Frage nach der Weiterempfehlungsbereitschaft im Hinblick auf die Universität Bonn in den Fragenkatalog mit aufgenommen, so dass derzeit Daten aus drei Messzeitpunkten vorliegen.

Auf einer Skala von 0 bis 10, wie wahrscheinlich ist es, dass Sie die Universität Bonn einem Freund oder Bekannten weiterempfehlen werden?

Die Ergebnisse aus den Befragungen zeigen, dass die Studierenden die Universität Bonn nicht uneingeschränkt weiterempfehlen. Im Wintersemester 2006/2007 haben sich 1322 Studierende an der Befragung beteiligt; davon haben 18,2% der Studierenden die Universität Bonn aktiv weiterempfohlen, 42,6% sind passiv zufrieden, 39,2% sind Kritiker. Dies bedeutet, dass der größte Anteil der Studierenden passiv zufrieden ist, demnach der Universität zwar nicht schadet, aber auch keine positive Mundpropaganda betreibt. Der große Anteil der Kritiker spiegelt die Unzufriedenheit der Studierenden mit der Universität wider, mehr noch: Kritiker sind so unzufrieden, dass sie ihr Umfeld auf ihre Unzufriedenheit aufmerksam machen und somit u. a. möglicherweise potenzielle Studienanfänger davon abhalten, an der Universität Bonn ein Studium aufzunehmen.

Im Vergleich zum Vorjahr könnte sich aus der Bewertung der Universität Bonn im Wintersemester 2007/2008 durch einen erhöhten Anteil der Kritiker (41,8%) und geringere Anteile von passiv Zufriedenen (40,3%) und Promotoren (17,8%) ein negativer Trend ableiten lassen, der sich schwerlich durch die leicht geringere Zahl der Befragungsteilnehmer (n=1250) erklären lässt. Im Wintersemester 2008/2009 nahmen 1324 Studierende an der Befragung teil. Der Net Promoter Score hat sich im Vergleich zum vorigen Jahr leicht verbessert, eine grundsätzliche Tendenz lässt sich allerdings nicht ableiten (s. Tabelle 1). Man darf also auf die Ergebnisse und Entwicklungen der nachfolgenden Jahre unter zwei Aspekten gespannt sein: Zum einen, ob die in vorherigen Beiträgen (Hörsch & Rudinger zur Qualitätssicherung von Lehre und Studium an Hochschulen; Costa zum Bonner Modell der Hochschulevaluation) schon erwähnten Maßnahmen auch hier zu extensiverer Beteiligung führen – nicht zuletzt, um das »Argument« zu entkräften, es handele sich um

(kleine) verzerrte Stichproben dominiert von Kritikern – , zum anderen, ob sich mittelfristig eine Trendumkehr ins Positive einstellt.

Nichtsdestoweniger sollte die Universität Bonn aufgrund dieser Ergebnisse schon jetzt Ursachenforschung betreiben, warum der Anteil der Kritiker so hoch ist, welche Schritte unternommen werden können, um die passiv Zufriedenen zu Promotoren zu machen und wie die bisherigen Promotoren als solche gehalten werden können.

Tabelle 1 Net Promoter Score in der allgemeinen Studierendenbefragung

	Wintersemester 2006/2007	Wintersemester 2007/2008	Wintersemester 2008/2009
<i>Kritiker</i>	39,2%	41,8%	41,5%
<i>Passiv Zufriedene</i>	42,6%	40,3%	38,1%
<i>Promotoren</i>	18,2%	17,8%	20,5%
<i>Net Promoter Score</i>	-21,0	-24,0	-21,0

Erkenntnisse, in welchen Bereichen der Universität Handlungsbedarf besteht, lassen sich aus den Ergebnissen der gesamten Studierendenbefragung zu einzelnen Bereichen des »Unternehmens« Universität ableiten. Um mit dem Positiven zu beginnen: So heben sich beispielsweise die Öffnungszeiten der Bibliothek, das Angebot des HRZ (Hochschulrechenzentrum, Zentraler IT-Service der Universität) und das Sportangebot der Universität durch eine positive Bewertung ab, sodass in diesen Bereichen kein Veränderungsbedarf zu bestehen scheint, sondern das derzeitige Angebot vielmehr aufrechterhalten bzw. erweitert werden sollte.

Andere Rahmenbedingungen des Studiums werden jedoch deutlich schlechter bewertet. So gibt z.B. die Mehrzahl der Studierenden an, »eher unzufrieden« mit der Ausstattung der Gebäude an der Universität, der Vermittlung von berufschancenverbessernden Kompetenzen in den Veranstaltungen und der Beteiligung Studierender an der Forschung zu sein. Unter Umständen begründet sich auch in diesen Punkten die niedrige Weiterempfehlungsbereitschaft. Mögliche daraus resultierende Maßnahmen wären ein verstärktes Angebot an Fortbildungen oder Trainings zum Erwerb von fächerübergreifenden Kompetenzen und Kenntnissen, die die Berufschancen verbessern (s. dazu inzwischen das Angebot der Zentralstelle für Schlüsselkompetenzen – ZSK), eine stärkere Einbindung der Studierenden in Forschungsprojekte sowie die Verbesserung der Ausstattung der Gebäude.

Lediglich 5,5 % der Studierenden sind der Überzeugung, Einflussmöglichkeiten auf die Entwicklung und das Angebot der Universität Bonn zu haben. Um die angestrebte Bindung der Studenten als »Klienten« an die Hochschule zu bewirken, gilt es auch hier, Veränderungen einzuleiten und die Studierenden besser über ihre Einflussmöglichkeiten durch Engagement in Gremien oder durch Evaluation von Modulen und Lehrveranstaltungen mit

daraus erwachsenden Konsequenzen zu informieren. Nicht zuletzt sind hier aber auch die Studierenden selbst mit ihren Vertretern in den diversen Gremien wie beispielsweise Fachschaften oder Studienbeitragskommissionen aufgerufen, ganz im Sinne von aktiven Klienten ihre Interessen zu vertreten und nicht in eine Kundenrolle im schlechtesten Sinne zu fallen.

Absolventenbefragung

Im Rahmen von Absolventenbefragungen dient der Net Promoter Score als Indikator nicht nur für die Bindung der Absolventen an »ihre« Hochschule, sondern auch für bestimmte Maßnahmen und Serviceeinrichtungen, die es zu schaffen bzw. zu steigern gilt und die die Bindung der Absolventen positiv beeinflussen können. Beispielsweise haben, wie Lenecke (2005) zeigen konnte, hochschulnahe Beratungs- und Betreuungsangebote (»Fachstudienberatung«, »Berufseinstiegsberatung«, »Betreuung bei Prüfungsangelegenheiten«) einen signifikanten Einfluss auf die Hochschulbindung.

Für eine wirkungsvolle Alumniarbeit sind Indikatoren wie der Net Promoter Score unverzichtbar, lassen sie doch Vergleiche der Bindungsqualität derselben Absolventen über die Zeit und verschiedener Absolventenkohorten untereinander zu. In Kombination mit der rückblickenden Bewertung von Studium und Lehre und des Berufseinstiegs lassen sich Stärken und Schwächen der Hochschulausbildung und Berufsvorbereitung durch die Universität erkennen. Der Net Promoter Score ist sozusagen ein Gesamtindikator für die Bindung an die Hochschule. Anhaltspunkte für Veränderungen ergeben sich aus den Befragungsergebnissen. Die Alumniarbeit verfolgt das Ziel, langfristige Bindung der Absolventen an die Universität zum Aufbau von Netzwerken zu schaffen, indem Belange und Kritik der Absolventen für qualitätssichernde Maßnahmen ernst genommen werden, denn Absolventen schicken möglicherweise die »Studierenden« von morgen an die Universität.

Bindung bewirkt man aber nicht nur durch Aktionen und *events* wie Begrüßungsfeiern, Absolventenfeste oder Alumni-Angebote, sondern in erster Linie dadurch, dass die Universität die Studierenden bereits während ihres Studiums oder besser noch gleich zu Studienbeginn ernst nimmt, sie als »Kunden« insofern ansieht, als dass sie für die Universität positive wie negative Wirkung haben können. Nur der zufriedene Student wird zum zufriedenen Absolventen und bleibt auch nach seinem Studienabschluss der Universität verbunden. Im Idealfall kann er vom Netzwerk der Universität profitieren, zugleich profitiert die Universität von ihm als Förderer.

Absolventinnen und Absolventen der Universität werden ein Jahr, fünf und zehn Jahre nach ihrem Studienabschluss an der Universität Bonn eingeladen, Lehre und Studium an der Universität Bonn rückblickend zu bewerten und über ihren Berufseinstieg Auskunft zu geben. Inhalte dieser Befragung sind Verlauf des Berufseinstiegs, Berufstätigkeit, Zusatzqualifikation, rück-

blickende Zufriedenheit mit Studieninhalten, universitäre Rahmenbedingungen, Bewertung der Forschungsmöglichkeiten, Image der Universität und die Weiterempfehlungsbereitschaft. Die hier vorgestellten Ergebnisse stammen aus der Befragung ein Jahr nach dem Studienabschluss, bislang liegen noch keine Ergebnisse für die Befragungen nach fünf bzw. zehn Jahren vor.

Insgesamt haben sich 1023 Absolventinnen und Absolventen hinsichtlich ihrer Weiterempfehlungsbereitschaft geäußert (Stand: Juli 2009). 23,9% der Absolventinnen und Absolventen der Universität Bonn betreiben positive Mundpropaganda für die Universität, 37,8% sind passiv zufrieden, 38,3% sind Kritiker, die ihr Missfallen offen kundtun (s. Tabelle 2). Unter dem Aspekt, dass Absolventenerfolgs- und -verbleibsstudien Gegenstand der Ziel- und Leistungsvereinbarungen III mit dem Lande Nordrhein-Westfalen sind, wird auch hier eine größere Beteiligung vor allem durch systematische Kooperation der Prüfungsämter erwartet.

Tabelle 2 Net Promoter Score in der Absolventenbefragung (Stand: Juli 2009)

<i>Kritiker</i>	38,1%
<i>Passiv Zufriedene</i>	38,3%
<i>Promotoren</i>	23,6%
<i>Net Promoter Score</i>	-14,5

Unabhängig davon sollte schon auf dieser Basis den Ursachen für den ja ebenfalls negativen Net Promoter Score nachgegangen werden. Sie liegen wohl vor allem in der mangelnden bzw. fehlenden Unterstützung der Universität oder des Fachs beim Übergang in den Beruf. 84% der Absolventen gaben an, dass es keine Form der Unterstützung gegeben habe, 89% erachten ein solches Angebot jedoch als wünschenswert.

Die Nutzung der vorhandenen Netzwerke, die Vermittlung von Kontakten zu Arbeitgebern oder Möglichkeiten für Bewerbertrainings sind nur Beispiele für mögliche Maßnahmen, die aus der Befragung resultieren. Die Universität Bonn hat im Jahre 2005 ein Career Center eingerichtet. Angeboten wird neben einer Jobbörse auch die qualifizierte Beratung und Unterstützung von Absolventen beim Übergang in den Beruf. Bisher hat sich das Angebot des Career Center noch nicht auf die Weiterempfehlungsbereitschaft der Absolventinnen und Absolventen ausgewirkt. Es wird sich in Zukunft zeigen, ob das Angebot des Career Centers langfristig auch den Net Promoter Score der gesamten Universität beeinflusst.

Ursachen für Unzufriedenheit finden sich auch in der mangelnden Beratung durch die Zentrale Studienberatung und die ungenügenden Beratungs- und Orientierungshilfen auf Fakultäts- und Lehrereinheitsebene. Auch hier hat die Universität Bonn bereits begonnen, gegenzusteuern. Studienanfänger und -innen erhalten mit dem *Studienkompass* einen »Wegweiser zum Studium in

Bonn«, in dem wichtige Informationen zu Beratung, Organisation des Studiums, Unterstützungsmöglichkeiten aber auch Tipps zur Freizeitgestaltung oder zur Studienfinanzierung enthalten sind. Der Effekt des *Studienkompass* wird sich wahrscheinlich in drei bis vier Jahren zeigen, wenn diese Studienanfänger und -innen ihr Studium abschließen.

Ein weiteres Serviceangebot für angehende Studierende an der Universität Bonn sind Online Self-Assessment-Verfahren (vgl. Sindern & Pietrangeli zur Entwicklung fachspezifischer Online Self-Assessments in diesem Band) und ein zielgruppenorientiertes Studierenden-Informationen-System. Inwieweit diese Angebote sich auf die Bindungsqualität der Studierenden an die Universität Bonn auswirken, wird sich ebenfalls erst im Verlauf der nächsten Jahre zeigen, denn – wie auch an anderen Stellen gesagt – solche Prozesse beanspruchen Zeit für eine tiefgreifende Implementierung.

Befragung der Verwaltungsmitarbeiter/innen

Die Befragung der Mitarbeiterinnen und Mitarbeiter der Universitätsverwaltung erfolgte im Zusammenhang mit der Einführung eines elektronischen Projektmanagementsystems über vier Befragungszeitpunkte (vgl. Ponert & Hörsch zur Evaluation von Change Management am Beispiel der Bonner Universitätsverwaltung in diesem Band). Ziel der Befragung war die Erfassung der Stärken, Schwächen und der Akzeptanz der Veränderung bei den Betroffenen. Inhalte der Befragung waren Bekanntheit des Projektmanagementsystems, Einsatz, Stärken und Schwächen des Systems sowie wahrgenommene Veränderungen. Die Befragung zu vier unterschiedlichen Zeitpunkten von Beginn bis zum Abschluss der Einführung des neuen Projektmanagementsystems sollte mögliche Veränderungen innerhalb der Einführungsphase aufzeigen. Tabelle 3 zeigt die vier Befragungszeitpunkte im Überblick.

Tabelle 3 Die vier Befragungszeiträume im Überblick

	<i>T-Null</i>	<i>T-Eins</i>	<i>T-Zwei</i>	<i>T-Drei</i>
<i>Größe der Stichprobe</i>	N=412			
<i>Befragungszeitraum</i>	02.08.- 20.08.2007	16.11.- 21.12.2007	17.03.- 21.03.2008	03.11.- 25.11.2008
<i>Rücklauf</i>	n ₀ =211 (52,49%)	n ₁ =118 (29,35%)	n ₂ =129 (32,08%)	n ₃ =174 (39,5%)

Zum jeweiligen Befragungsabschluss wurden alle kontaktierten Mitarbeiterinnen und Mitarbeiter (n=412) gebeten, ihre Weiterempfehlungsbereitschaft im Hinblick auf die Universitätsverwaltung als Arbeitgeber sowie im Hinblick auf die Universität als Hochschule einzuschätzen:

*Auf einer Skala von 0 bis 10, wie wahrscheinlich ist es, dass Sie die Verwaltung der Universität Bonn einem Freund oder Bekannten als **Arbeitgeber** weiterempfehlen?*

*Auf einer Skala von 0 bis 10, wie wahrscheinlich ist es, dass Sie die Universität Bonn als **Hochschule** einem Freund oder Bekannten weiterempfehlen?*

Die Antworten auf beide Weiterempfehlungsfragen lassen über die vier Befragungszeitpunkte bei deutlichen Schwankungen doch mehr oder minder klare Abwärtstrends erkennen (s. für die erste Weiterempfehlungsfrage Tabelle 4). Die Frage zur Weiterempfehlung als Arbeitgeber muss aber vor allem im Kontext des Veränderungsprozesses durch die Einführung des Projektmanagementsystems betrachtet werden: In die Bewertung der Universitätsverwaltung als Arbeitgeber ist möglicherweise der Unmut über die Einführung oder Schwierigkeiten in der Anwendung des Projektmanagementsystems eingeflossen. Interessant ist hier insbesondere der Anteil der passiv Zufriedenen, der sich zum letzten Befragungszeitpunkt T-Drei auf dem Höchststand von 43,7% befindet. Dies wirft ein durchaus interessantes Licht auf solche Befragungen, die schlicht Zufriedenheit (mit der Tätigkeit, dem Arbeitgeber etc.) ermitteln, da die »Färbung« der Zufriedenheit (aktiv/passiv) nicht deutlich wird. Es gilt aber, den Anteil der passiv Zufriedenen, die sich eben zufriedenstellend arrangiert haben, dadurch zu verringern, dass diese passiv zufriedenen Mitarbeiter zu Förderern gemacht werden, die die Universitätsverwaltung als Arbeitgeber aktiv weiterempfehlen und somit ihrer Loyalität gegenüber ihrem Arbeitgeber Ausdruck verleihen. Mögliche Maßnahmen sind an dieser Stelle beispielsweise Unterstützung von Mitarbeitern

durch Fortbildungsmöglichkeiten, Förderung der persönlichen Weiterentwicklung durch gemeinsame Zielvereinbarungen und Interesse an den Belangen der Mitarbeiter durch Ansprechbarkeit der Vorgesetzten zu zeigen.

Tabelle 4 Net Promoter Score der Universitätsverwaltung als Arbeitgeber in der Mitarbeiterbefragung

	<i>T-Null</i>	<i>T-Eins</i>	<i>T-Zwei</i>	<i>T-Drei</i>
<i>Kritiker</i>	41,1%	39,8%	44,4%	41,4%
<i>Passiv Zufriedene</i>	35,6%	36,5%	30,6%	43,7%
<i>Promotoren</i>	23,3%	23,7%	25,0%	14,9%
<i>Net Promoter Score</i>	-17,8	-16,1	-19,4	-26,4

In der ersten Befragung (T-Null) unmittelbar nach Beginn des Einführungsprozesses haben sich 209 Mitarbeiterinnen und Mitarbeiter hinsichtlich ihrer Bereitschaft, die Verwaltung der Universität als Arbeitgeber weiter zu empfehlen, geäußert. Zum Zeitpunkt T-Eins waren es 114, bei T-Zwei 131 und bei der letzten Befragung (T-Drei) 157, deutlich höhere Anteile als bei den vorherigen Befragungen. Hier von selbstselektiven Stichproben auszugehen, bei der sich viermal gerade nur die Promotoren zurückgehalten haben, wäre eine etwas gewagte Annahme, wie das folgende Ergebnis veranschaulichen mag.

Insgesamt stellt sich der Net Promoter Score für die Universität Bonn als Organisation in Bezug auf ihre zentralen Aufgaben und ihr Selbstverständnis als Universität nämlich deutlich positiver dar. Die unterschiedliche Wahrnehmung und Bewertung der Universität Bonn als Arbeitgeber auf der einen und als Hochschule auf der anderen Seite resultiert wahrscheinlich aus der situationsbedingten Unzufriedenheit der Verwaltungsmitarbeiter aufgrund von Umstrukturierungen und Neuerungen, die z.T. erheblichen Mehraufwand der Mitarbeiter fordern, während Identifikation mit der Universität durchaus vorhanden zu sein scheint.

Aber auch diese Bewertung der Universität als Hochschule ist sehr starken Schwankungen ausgesetzt (s. Tabelle 4). Möglicherweise gehen diese Schwankungen auch darauf zurück, dass nicht zu jedem Zeitpunkt dieselben Mitarbeiter teilgenommen haben. Darüber wurden keine Informationen erhoben. Offenbar ist jedoch die Identifikation der Befragten nicht ausreichend gefestigt, das zeigt sich insbesondere in dem großen Anteil passiv zufriedener Mitarbeiter.

Tabelle 5 Net Promoter Score der Hochschule in der Mitarbeiterbefragung

	<i>T-Null</i>	<i>T-Eins</i>	<i>T-Zwei</i>	<i>T-Drei</i>
<i>Kritiker</i>	26,1	29,7	31,6	29,3
<i>Passiv Zufriedene</i>	38,7	37,2	29,6	44,3
<i>Promotoren</i>	35,2	33,1	38,8	26,4
<i>Net Promoter Score</i>	9,1	3,4	7,2	-2,8

Um noch einmal die Stichproben- bzw. Methodendiskussion aufzugreifen: Eine klare Beziehung zwischen Stichprobengröße und Net Promotor Score lässt sich hier nicht ausmachen, der zeitliche (negative) Trend über die vier Erhebungen ist nicht gar so eindeutig ausgeprägt.

Die Erwartungen der befragten Mitarbeiter an die Universität Bonn sind sicherlich nicht in dem Ausmaß erfüllt, wie es bei einem erfolgreichen »Unternehmen« angezeigt wäre. Durch verstärkte Kommunikation zwischen Führungskräften und Mitarbeitern, aber auch durch den Dialog der Mitarbeiter untereinander können Angebote wie Weiterbildungsmöglichkeiten, Stärken und Erfolge innerhalb der Universität verbreitet werden. Zudem gilt es, denen, die Verantwortung tragen, entsprechende Anerkennung entgegen zu bringen und die Mitarbeiter als starke Gemeinschaft und zugleich Teil der Universität zu motivieren. Dabei ist es wichtig, mit den vorhandenen aktiven Kanälen und einem möglichst großen »Face-to-Face«-Anteil zu arbeiten, um das Vertrauen und die Motivation der Mitarbeiter zu erhalten und zu stärken. Das gilt auch für »schlechte Nachrichten«, die ehrlich und offen mitgeteilt werden sollten.

Fazit

Vergleiche mit anderen Hochschulen sind aufgrund des seltenen Einsatzes des Net Promotor Scores an Universitäten nicht durchführbar, so dass erst einmal »nur« interne bzw. zeitliche Vergleiche, d. h. Betrachtung der Veränderungen des Indexes möglich sind.

Das vorläufige Fazit baut auf den bisherigen Untersuchungen mit ihren Beschränkungen auf. Für die Fortschreibung dieser Untersuchungen sind wachsende Stichprobenumfänge zumindest bei den Studierenden- und Absolventenbefragungen wünschenswert (zu Möglichkeiten der Steigerung der Teilnahmebereitschaft vgl. Hörsch & Rudinger zur Qualitätssicherung von Lehre und Studium an Hochschulen in diesem Band). Durch die kontinuierliche Erfassung des Net Promoter Scores unter Beteiligung der Studierenden und Absolventen lassen sich dauerhafte Entwicklungen aufzeigen. Hier gilt es, von allen Seiten aktiver zur Teilnahme zu motivieren und die Einflussmöglichkeiten auf die Prozesse innerhalb der eigenen Universität deutlicher zu machen.

In allen Befragungen zeigt sich, dass der größte Anteil bei Studierenden, Absolventen und Mitarbeitern im Bereich der passiv Zufriedenen angesiedelt ist. Daraus lässt sich ableiten, dass die Universitätsangehörigen grundsätzlich zufrieden sind, wie auch entsprechende allgemeine Zufriedenheitsbefragungen zeigen, aber die Universität Bonn eben nicht aktiv weiterempfehlen. Darin lässt sich großes Potenzial erkennen, die passiv Zufriedenen langfristig zu Förderern der Universität zu machen. Entsprechende Maßnahmen zur Stärkung der Bindung lassen sich aus den einzelnen Befragungskontexten ableiten.

Gerade für die Universität ist es wichtig, dass sich alle Akteure mit der Universität sozusagen aktiv verbunden fühlen. Ob Mitarbeiter, Student, Absolvent oder gar schon Schüler (zumindest aus der Region): Durch verschiedene universitäre Maßnahmen und Aktionen sollen die Gemeinschaft und die Identifikation mit der Universität Bonn gefördert werden. Die ganzheitliche Betrachtung des Systems »Universität« ist unabdingbar, um in der deutschen Hochschullandschaft auch zukünftig konkurrenzfähig zu bleiben. In diesem Sinne verfolgt die Universität Bonn das Ziel, die schon zahlreich existierenden Einzelmaßnahmen in ein Gesamtkonzept zu integrieren, um dieses öffentlichkeitswirksam im Wettbewerb sichtbar machen zu können. Die Universität Bonn baut bei der Studierendenbindung auf Service, Tradition, Organisation, Gemeinschaft und gute Lehre. Da Bindungsmanagement bereits vor der Entscheidung zum Studium an der Universität Bonn ansetzen kann, gibt es verschiedene Initiativen wie Kinderuni, Förderprogramme für begabte Schüler sowie Online-Self-Assessment-Verfahren, die maßgeblich zur Profilbildung und Attraktivität der Universität Bonn beitragen.

Ziel der Universität sollte es sein, Studierende zu erfolgreichen Alumni zu machen und Alumni zu Freunden und Förderern der Universität. Aus der nachhaltigen Bindung insbesondere von Studierenden und einem attraktiven Alumniangebot kann es vom Friendraising langfristig zum Fundraising kommen. Dabei steht die Beziehung zwischen Alumna/Alumnus und Universität im Vordergrund, die für beide Seiten z.B. in Form von Dienstleistungen von Nutzen ist. Unternehmer können beispielsweise in die Universität »investieren« und zugleich erfolgreiche Studierende als Praktikanten oder zukünftige Mitarbeiter rekrutieren. Voraussetzungen für Friendraising sind neben dem attraktiven Alumniangebot eine enge Bindung an die Universität sowie Zufriedenheit und Dankbarkeit in Verbindung mit der Ausbildung und Förderung an der Universität (Giebisch & Langer, 2005).

Einige Hochschulen sind bereits auf dem besten Wege, Initiativen in dieser Richtung aufzubauen (Giebisch & Langer, 2005). An der Universität Passau werden über den Career-Service qualifizierte Absolventen und Unternehmer zusammengebracht, eine jährlich stattfindende Firmenkontaktmesse bietet die Möglichkeit für Absolventen und angehende Absolventen, mit Unternehmen bereits erste Kontakte zu knüpfen und umgekehrt. Durch den

ständigen Dialog zwischen Hochschule und Wirtschaft werden Partnerschaft und Praxisnähe als Qualitätsmerkmal betrachtet und gelebt (Dirscher, 2008).

Der Einfluss der Bindung auf das monetäre Wachstum (Keiningham et al., 2007) lässt sich in Unternehmen prüfen. Analog zum monetären Wachstum ist an Hochschulen Attraktivität und Image entscheidend, messbar z.B. in Studienbewerberzahlen, Studierendenzahlen, Absolventenzahlen, Studienabbrecherquoten, Mitarbeiterschwund. Die Verknüpfung dieser Monitoringdaten mit den durch Befragungen erhobenen Stärken und Schwächen der Universität Bonn bringt die Erkenntnis über den Zusammenhang von Anliegen und Bedürfnissen der Universitätsangehörigen einerseits und Kennziffern und deren Entwicklungspotenzial andererseits. In diesem Zusammenhang ist die wesentliche Voraussetzung für die Akzeptanz von Notwendigkeit organisationalen Wandels die größtmögliche Transparenz von Prozessen und Entscheidungen. Zugleich gilt es, für den Veränderungsprozess ausreichend Zeit einzuräumen.

Literatur

- DIRSCHER, K. (2008). Fundraising beginnt mit Friendraising – Das Modell des Neuberger Gesprächskreises an der Universität Passau. *Beiträge zur Hochschulforschung*, Heft 3, 30. Jahrgang, 2008. S. 92- 104. Zugriff am 28.08.2009 unter http://bildungsklick.de/datei-archiv/50640/fundraising-3-2008_gesamt-1.pdf
- GIEBISCH, P. & Langer, M. (2005). *Erste Eindrücke zum Stand des Hochschulfundraising in Deutschland*. (Centrum für Hochschulentwicklung (CHE) Arbeitspapier Nr. 68, Dezember 2005). Zugriff am 28.08.2009 unter http://www.che.de/downloads/CHE_Fundraising_AP68.pdf
- KEININGHAM, T.L., Cooil, B., Andreassen, T.W. & Aksoy, L. (2007). A longitudinal examination of Net Promoter and firm revenue growth. *Journal of Marketing*, Vol 71, 39-51.
- LANGER, M., Ziegele, F. & Hennig-Thurau, T. (2001). *Hochschulbindung – Entwicklung eines theoretischen Modells, empirische Überprüfung und Ableitung von Handlungsempfehlungen für die Hochschulpraxis*. Gütersloh: Centrum für Hochschulentwicklung. Zugriff am 28.08.2009 unter www.che.de/downloads/hs_bindung_bericht0105.pdf
- LENECKE, K. (2005). *Hochschulbindung durch Student Services. Grundlagen, Analysen, Perspektiven*. Saarbrücken: Dr. Müller.
- REICHHELD, F.F. (2003,). The one number you need to grow. *Harvard Business Review*, December 1st, 3-11.
- REICHHELD, F.F. (2006). *The ultimate question*. Boston: Harvard Business School Press.
- REICHHELD, F.F. & Seidensticker, F.-J. (2006). *Die ultimative Frage*. München: Hanser Verlag.

SCHWAIGER, M. (2003). Der Student als Kunde – eine empirische Analyse der Zufriedenheit Münchner BWL-Studenten mit ihrem Studium. *Beiträge zur Hochschulforschung*, 25 (1), 32-62. Zugriff am 28.08.2009 unter http://www.ihf.bayern.de/beitraege/2003_1/1-2003%20Schwaiger.pdf

Alexandra Borsch

Der Net Promoter Score (NPS¹) als diagnostisches Instrument zur Analyse der Loyalität interner Kunden

Die Loyalität externer Kunden wird gemein hin als bedeutender Faktor langfristigen Unternehmenserfolgs anerkannt und hat sich in Form des NPS als Wachstumsprädiktor erwiesen. Gegenstand des vorliegenden Beitrags ist die Anwendung des NPS auf den Bereich interner Kunden im Unternehmens- und Hochschulkontext sowie eine quantitative Untersuchung der zugrunde liegenden Determinanten der Weiterempfehlungsbereitschaft. Hypothesenkonform erwiesen sich die Aspekte Betriebs-/Hochschulklima, Vertrauen, Image und affektives Commitment als Treiber des NPS sowie das Vorhandensein zusätzlicher Angebote, die eine organisationale bzw. akademische Integration der Mitarbeiter und Studenten fördern. Personale Variablen wie der soziodemographische Hintergrund und die Persönlichkeit der internen Kunden hingenen standen in keinem Zusammenhang zur ihrer Loyalität. Ausgehend von dieser empirischen Basis werden mögliche Handlungsimplikationen diskutiert.

Einleitung

Im innerbetrieblichen Bereich gewinnt der Wert der Loyalität, operationalisiert über die Weiterempfehlungsbereitschaft (Mund-zu-Mund-Propaganda), zunehmend an Bedeutung. So resultieren die zum Erhalt der Wettbewerbsfähigkeit vorgenommenen organisationalen Umgestaltungen in Strukturen, die sich durch Flexibilität und Veränderungsbereitschaft auszeichnen. Derartige Tugenden werden zunehmend auch von Unternehmensangehörigen gefordert (Gomez, 2000). Insbesondere viel versprechende Fach- und Führungsnachwuchskräfte wechseln in Folge innerhalb ihrer ersten zehn Berufsjahre bis zu fünf Mal ihre Arbeitsstelle, denn eine ununterbrochene Beschäftigung bei einem Unternehmen gilt als Anzeichen für beschränkte Erfahrungen und einen Mangel an Ehrgeiz (Gomez, ebd.). Verstärkt wird der Drang der Erwerbstätigen nach »Career-Surfing« und »Job-Hopping« durch den so genannten »War for Talent«, der eine Situation auf dem Arbeitsmarkt widerspiegelt, die durch starke Bemühungen von Unternehmen gekennzeichnet

1 Net Promoter is a registered trademark of Satmetrix Systems, Inc., Bain & Company and Fred Reichheld.

net ist, die talentiertesten Mitarbeiter zu finden (Chambers, Foulon, Handfield-Jones, Hankin & Michaels, 1998).

Solange zum einen Flexibilität und Dynamik Positivmerkmale im Lebenslauf darstellen und zum anderen die Arbeitsplatzsicherheit von Unternehmensseite nicht garantiert werden kann, wird zunehmend eine Form der Loyalität wichtig, die über eine Anstellung hinaus ihre Wirkung erzielt. Folglich sollte die Weiterempfehlungsbereitschaft der Mitarbeiter fokussiert werden, da diese Art außerordentlichen Engagements gegenüber der Organisation sowohl die Rekrutierung potenzieller talentierter Mitarbeiter als auch die Neukundengewinnung positiv beeinflusst (IAB, 2006; Scherrer, 1975).

Auch das *Hochschulwesen* in Deutschland unterliegt gegenwärtig diversen Veränderungen, welche die Profilbildung der Universitäten und den Wettbewerb um qualifizierte, zahlungsfähige Studenten notwendig machen (Rindermann & Oubaid, 1999). Wie im Falle der Mitarbeiter gilt auch bei den Studierenden die Maxime, jeder Student müsse zumindest einmal im Laufe seines Studiums die Hochschule wechseln, um seine Kompetenzen zu erweitern (Langer, Ziegele & Hennig-Thurau, 2001). Aus diesem Grund erscheint es relevant, neben der Loyalität der Mitarbeiter auch die Weiterempfehlungsbereitschaft von Studenten in Bezug auf ihre Hochschulen mithilfe des NPS zu erheben und sie auf ihre Determinanten hin zu untersuchen.

Beim Net Promoter Score handelt es sich ursprünglich um eine Kenngröße der Kundenloyalität. Reichheld operationalisiert dabei mit einer einzigen Frage die Weiterempfehlungsbereitschaft externer Kunden, die er als Facette der Loyalität ansieht. Kunden, die Leistungen und Produkte in ihrem Freundeskreis weiterempfehlen, betreiben kostenlose Mund-zu-Mund-Propaganda und engagieren sich auf diese Weise für das jeweilige Unternehmen (Reichheld, 2003). Der NPS stellt dabei das Resultat einer Befragung dar, bei der ausschließlich die folgende Frage gestellt wird: »Wie wahrscheinlich ist es, dass Sie das Unternehmen XY an einen Freund oder Kollegen weiterempfehlen?«. Beantwortet wird diese Frage auf einer Skala von 0 bis 10, wobei die 0 bedeutet, dass das Unternehmen ganz bestimmt nicht und die 10, dass das Unternehmen ganz bestimmt weiterempfohlen wird.

Nach Reichheld (2003) lassen sich die Befragten in Abhängigkeit ihrer Antworten in drei Gruppen einteilen: Promoters, Passives und Detractors.

- Unter *Promotern* (Förderern) werden Personen verstanden, die das Unternehmen enthusiastisch mit einer hohen Wahrscheinlichkeit weiterempfehlen (Ratings 9 und 10)
- Als *Passives* werden hingegen Kunden bezeichnet, die lediglich passiv zufrieden sind und das Unternehmen nur zum Teil weiterempfehlen (Ratings 7 und 8).
- Bei *Detractors* (*Kritiker*) handelt es sich um enttäuschte Kunden, die das Unternehmen mit einer geringen Wahrscheinlichkeit weiterempfehlen (Ratings 0 bis 6)

Der Net Promoter Score kann Werte zwischen plus und minus 100% annehmen und wird aus den Antworten der Kunden auf die oben genannte Frage berechnet, indem der Prozentsatz der Kritiker vom Prozentsatz der Promoter subtrahiert wird. Auf diese Weise erhält man nach Reichheld (2003) ein Maß für die Kundenloyalität, welches aufgrund seiner einfachen und zeitsparenden Ermittlung ökonomisch ist und effektiv genutzt werden kann. Dieses Konzept der Kundenloyalität wird nun im Kontext interner Kunden untersucht.

Theoretische Fragestellungen und Hypothesen

Das Anliegen dieses Beitrags besteht in der statistischen Identifikation möglicher Treiber des NPS. Dabei wird zwischen zwei verschiedenen Fragestellungen unterschieden:

- (1) Inwiefern eignen sich die potenziellen Loyalitätstreiber als Korrelate der Weiterempfehlungsbereitschaft?
- (2) Unterscheiden sich die verschiedenen nach Reichheld (2003) konzeptualisierten Loyalitätsgruppen Detractors, Passives und Promoters in ihren Ausprägungen bzgl. der unabhängigen Variablen? Bzw. anhand welcher Variablen können die verschiedenen Loyalitätsgruppen differenziert werden?

Zur Erfassung möglicher Loyalitätsdeterminanten wurden in einem ersten Schritt 24 Mitarbeiter und Studenten in halbstrukturierten Interviews zur ihrer Weiterempfehlungsbereitschaft sowie ihren zugrunde liegenden Motiven befragt. Basierend auf diesen Ergebnissen sowie theoretischen Vorüberlegungen wurden im Anschluss Forschungshypothesen formuliert und mithilfe eines quantitativen Online-Fragebogens untersucht. Dabei wurden die folgenden unabhängigen Variablen berücksichtigt.

Soziodemographische Aspekte

Vor dem Hintergrund von Untersuchungen aus dem Bereich der Mundpropaganda- (Ganesh, Arnold & Reynolds, 2000) und Hochschulforschung (Schwaiger, 2003), die aufzeigen konnten, dass insbesondere Neukunden bzw. Erstsemester eine positive Einstellung gegenüber ihrer Institution haben und sich aktiv für diese engagieren, wurde angenommen, dass die Dauer der Tätigkeit (Jobalter) sowie das Alter der Berufstätigen bzw. die Studiendauer Treiber der Weiterempfehlungsbereitschaft darstellen.

Persönlichkeit der Befragten

Da sowohl eine Metaanalyse von Judge, Heller und Mount (2002) aufzeigen konnte, dass die Arbeitszufriedenheit mit der Persönlichkeit zusammenhängt, als auch Sundaram, Mitra und Webster (1998) signifikante Korrelatio-

nen zwischen der Mundpropaganda externer Kunden und ihrer Persönlichkeit nachweisen konnten, wurde postuliert, dass die Persönlichkeit der Befragten ein Korrelat der Weiterempfehlungsbereitschaft darstellt. Insbesondere wurde davon ausgegangen, dass Extraversion, Verträglichkeit, Gewissenhaftigkeit sowie das soziale Anerkennungsbedürfnis positiv, die Persönlichkeitseigenschaften Offenheit und Risikomeidung jedoch negativ mit der Weiterempfehlung zusammenhängen.

Arbeits-/Studienzufriedenheit

Vor dem Hintergrund der Ergebnisse zahlreicher empirischer Studien, die im Bereich der Kundenloyalität eine Verbindung zwischen Zufriedenheit und Mundpropaganda aufweisen konnten (Giering, 2000), galt es zu prüfen ob derartige Zusammenhänge im Bereich der NPS-Forschung repliziert werden können. In Bezug auf Reichhelds (2006) empirisch begründete Annahme, ein hohes Maß an Zufriedenheit verursache nicht zwangsweise Loyalität, sollte ferner untersucht werden, ob sich die auf der Loyalitätsdimension »benachbarten« Passives und Promoters in ihrer Zufriedenheit unterscheiden.

Betriebs- /Hochschulklima

Basierend auf den empirischen Erkenntnissen von Organ (1988), die eine Erhöhung eines übervertraglichen freiwilligen Engagements von Arbeitnehmern auf eine Verbesserung des Organisationsklimas zurückführen konnten, wurde das wahrgenommene Betriebs-/Hochschulklima als Treiber des NPS in Betracht gezogen.

Affektives, normatives und kalkulatives Commitment

Da Felfe und Six (2006) das organisationale Commitment von Arbeitnehmern mit deren Engagement in Zusammenhang bringen konnten, wurde vermutet, dass auch im Bereich der Weiterempfehlungsbereitschaft das normative² und das affektive³ Commitment der Befragten als Treiber in Frage kommen, während das kalkulative Commitment, das Vorhandensein von Wechselbarrieren widerspiegelnd, zwar die Bindung, aber nicht die Loyalität erhöhen sollte.

2 Das normative Commitment spiegelt eine Verbundenheit wider, die auf einer Verinnerlichung von Normen beruht.

3 Das affektive Commitment beruht auf einer emotionalen Verbundenheit, die sich in gemeinsamen Werten und Stolz ausdrückt.

Vertrauen

Im Bereich externer Kunden konnte u.a. von Giering (2000) ein positiver Zusammenhang von Vertrauen und positiver Mundpropaganda aufgezeigt werden. Dieser Befund wurde von Graeff (1998) auch im Unternehmenskontext bestätigt. Demzufolge beeinflusst das Vertrauen – insbesondere zur Organisation – die Bindung der Mitarbeiter positiv. Vor diesem Hintergrund galt es zu prüfen ob sich diese Befunde im Bereich des NPS bestätigen lassen.

Fluktuationsabsicht

Die Fluktuations- bzw. Abwanderungsintention stellt ein Konstrukt dar, welches im Bereich der Bindungsforschung bei externen Kunden sowie Mitarbeitern und Studenten bereits häufig untersucht wurde und bisweilen auch mit Einstellungen korreliert (Gebert & von Rosenstiel, 2002). Da dieser Ansatz jedoch intervenierende Rahmenbedingungen unberücksichtigt lässt (Lenecke, 2005; Withey & Cooper, 1992) und weder Bindung noch Fluktuation erschöpfend mit einer positiven bzw. negativen Einstellung erklärt werden können, wurde unterstellt, die Fluktuationsabsicht eigne sich nur bedingt als Treiber der Loyalität.

Unternehmens-/Hochschulimage

In Bezug auf die Arbeit von Haase (1997), die den Wert wahrgenommener Unternehmensimages im Mitarbeiterbereich betont sowie der Untersuchung von Nguyen und LeBlanc (2001), die die Bedeutung des Hochschulimages für die Mundpropaganda von Studenten hervorhebt, wurde vermutet, dass von internen Kunden wahrgenommene institutionale Image stelle einen Treiber der Weiterempfehlungsbereitschaft dar.

Offerierte zusätzliche Angebote

Basierend auf den Ergebnissen von Tinto (1975) sowie Langer et al. (2001), die eine soziale und organisationale Integration als Determinante der Loyalität herausarbeiten konnten, wurde angenommen, dass zusätzliche Angebote von Institutionen die Weiterempfehlungsbereitschaft interner Kunden erhöhen. Demnach sollten sich die Befragten, deren Institutionen über zusätzliche Angebote verfügen, hinsichtlich der Weiterempfehlungsbereitschaft von denen unterscheiden, deren Institutionen keine Angebote offerieren.

Methoden

Die quantitative Befragung der Studierenden und Berufstätigen erfolgte webbasiert. Dazu wurde der Fragebogen in zwei unterschiedlichen Versionen

online programmiert und den Zielpersonen durch ein E-Mail-Schneeballsystem zur Verfügung gestellt.

Das Befragungsinstrument

Zur Erfassung der *Loyalität* der Befragten wurde der NPS nach Reichheld (2003) eingesetzt. Dabei wurde zwischen drei verschiedenen Operationalisierungen unterschieden. So wurden die Erwerbstätigen zum einen hinsichtlich ihrer Bereitschaft befragt, ihr Unternehmen als Arbeitgeber weiterzuempfehlen sowie zum anderen als Produzent/Dienstleister, während bei der Studierendenchprobe die Bereitschaft erhoben wurde, die jeweilige Hochschule weiterzuempfehlen.

Zur Beschreibung der Stichprobe und zum Zwecke weiterer Analysen wurden *soziodemographische Aspekte* der Befragten erfasst. So wurden die Probanden gebeten, ihr Alter, ihr Geschlecht und ihren höchsten erreichten akademischen Abschluss anzugeben. Darüber hinaus wurden die Erwerbstätigen zu folgenden zusätzlichen Angaben angehalten: Dauer der bisherigen Berufstätigkeit, monatliches Netto-Einkommen, durchschnittliche Wochenarbeitszeit sowie beruflicher Status.

Die Erfassung der *Persönlichkeit* der internen Kunden wurde mithilfe von 14 Aussagen realisiert, die sieben Persönlichkeitseigenschaften zugeordnet werden können. Dabei wurde jede Persönlichkeitsvariable durch zwei Items abgebildet, die auf einer fünfstufigen Likert-Skala von *trifft zu* bis *trifft nicht zu* beantwortet wurden. Die Formulierungen der Items wurden zwei verschiedenen Persönlichkeitsinventaren entliehen. Zur Erfassung der Dimensionen Neurotizismus, Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit wurde die deutsche *10-Item Version des Big Five Inventory* (BFI-10) von Rammstedt und John (2006) eingesetzt, während die Items zur Erfassung der Risikomeidung und des sozialen Anerkennungsbedürfnisses der deutschen Fassung des *Personality Research Form* (PRF) von Stumpf, Angleitner, Wieck, Jackson und Beloch-Till (1985) entliehen wurden.

Zur Prüfung des Zusammenhangs von *Zufriedenheit* und *Loyalität* wurde zum einen die Zufriedenheit mit der allgemeinen Arbeits-/Studiensituation erhoben. Die gleichzeitig dargebotene Antwortskala wurde in Anlehnung an diejenige des NPS entwickelt, um auf diese Weise eine vergleichende Prüfung beider Konstrukte zu ermöglichen. Ergänzend dazu wurde zum anderen die Zufriedenheit mit den folgenden spezifischen Aspekten der Institution erhoben: finanzielle Kompensation/Studiengebühren, Sicherheit des Arbeitsplatzes/Attraktivität des Hochschulstandortes, kollegiales Verhältnis, Reputation des Unternehmens/der Hochschule, Selbständigkeit, Lern- und Entwicklungsmöglichkeiten und Anforderungsvielfalt.

In einem umfangreichen Fragenblock mit 45 Items wurde das *wahrgenommene Betriebs-/Hochschulklima* erfragt. Zu diesem Zweck wurde eine

modifizierte Version des *Erhebungsbogens zur Erfassung des Betriebs- und Organisationsklimas* von von Rosenstiel (1992) eingesetzt und ferner an den Hochschulkontext angepasst. Die Subdimensionen des Betriebs-/Hochschulklimas sind der folgenden Tabelle 1 zu entnehmen.

Tabelle 1: Dimensionen der modifizierten Skala Betriebs-/Hochschulklima

<i>Betriebsklima-Dimensionen</i>	<i>Hochschulklima-Dimensionen</i>
Mitarbeiterorientierung	Studierendenorientierung
Unternehmenspolitik	Hochschulpolitik
Umgang unter den Kollegen	Umgang unter den Kommilitonen
Vorgesetztenverhalten	Lehrpersonal
Karriereförderung	Lehrangebot/Organisation des Studiums
Entlohnung	Information und Mitsprache
Information und Mitsprache	
Arbeitszeitregelung	
Arbeitsinhalte	
Kundenorientierung	

Zur Erfassung des organisationalen Commitments der internen Kunden wurde eine reduzierte Version des Fragebogens zur Erfassung von affektivem, kalkulatorischem und normativem Commitment gegenüber der Organisation, dem Beruf/der Tätigkeit und der Beschäftigungsform (COBB) von Felfe, Six und Schmook (2002) eingesetzt. Die 14 präsentierten Items bezogen sich dabei ausschließlich auf die Verbundenheit und Identifikation mit der Organisation. Die Beantwortung der Fragen erfolgte auf einer vierstufigen Skala von trifft zu bis trifft nicht zu.

Das Vertrauen gegenüber der Institution wurde anhand von vier Items gemessen, die den Skalen »Vertrauen zum Vorgesetzten« und »Vertrauen in das Unternehmen« von Graeff (1998) zugeordnet werden können. Diese Items wurden im Studierendenfragebogen an den Hochschulkontext angepasst und mit einer vierstufigen Antwortskala dargeboten.

Die zur Erfassung der Fluktuationsabsicht verwendeten Items wurden Baillo und Semmer (1994) entnommen und erforderten eine Einschätzung der »Wahrscheinlichkeit, in zwei Jahren noch bei demselben Unternehmen zu arbeiten, Gedanken an einen Stellenwechsel sowie die Bewerbungsabsicht bei einer attraktiven Stelle« (S. 155). Zur Beantwortung der Fragen wurde eine vierstufige verbal verankerte Skala dargeboten.

In einem weiteren Teil der Erhebung wurde das 1975 von Osgood, Suci und Tannenbaum entwickelte semantische Differenzial zur Messung des Unternehmens-/Hochschulimages verwendet. Bei diesem Verfahren wurden den Befragten semantische Skalen in Form bipolarer Eigenschaftswörter

vorgegeben, mit denen ihr Unternehmen/ihre Hochschule beurteilt werden sollten. Die Liste der 20 Wortpaare wurde mit einer sechsstufigen Rating-Skala verbunden, um die Stärke der Assoziationen und somit die affektive Bewertung des zu beurteilenden Konzeptes messen zu können. Die zu beurteilenden Wortpaare entstammten den Dimensionen Evaluation, Aktivität und Potenz. Zur Ergänzung des semantischen Differenzials wurde auf qualitativer Ebene eine freie Assoziation mit der jeweiligen Institution erfasst. So wurden die internen Kunden aufgefordert, die Frage »Wenn Ihr Unternehmen/Ihre Hochschule ein Tier wäre, um welches würde es sich dann handeln?« offen zu beantworten.

Zur Erfassung der sozialen bzw. organisationalen/akademischen Integration der Befragten wurde in Anlehnung an Tinto (1975) das Vorhandensein von Freizeit-, Dienstleistungs- und beruflichen/akademischen Angeboten erhoben. Im Falle der Erwerbstätigen wurden dabei 12 Angebote abgefragt, bei den Studierenden 24.

Stichprobenbeschreibung

An der quantitativen Mitarbeiterbefragung nahmen 165 Berufstätige teil. Von diesen Untersuchungsteilnehmern waren 87 weiblichen (52.7%) und 78 männlichen Geschlechts (47.3%), wobei das Durchschnittsalter 33.85 Jahre betrug ($SD = 10.16$). Der realisierte Fragebogen wurde ferner von 189 Studierenden ausgefüllt. Auch bei den Studierenden war die Geschlechterverteilung dabei mit 109 (58%) weiblichen und 80 (42%) männlichen Befragten relativ ausgewogen. Die Gruppe der Studierenden war zum Zeitpunkt der Befragung im Durchschnitt 25.16 ($SD = 3.72$) Jahre alt.

Ergebnisse

Die Darstellung der Ergebnisse gliedert sich nach den zwei verfolgten Fragestellungen, wobei zunächst die Ergebnisse korrelativer Analysen vorgestellt werden, bevor auf die Befunde in Bezug auf die postulierten Mittelwertsunterschiede in Abhängigkeit der Gruppenzugehörigkeit eingegangen wird, die mittels t-Tests und Varianzanalysen (ANOVA) inferenzstatistisch untersucht wurden.

Bedeutende korrelative Treiber

Bezüglich der Annahme *interindividueller Unterschiede* der Weiterempfehlungsbereitschaft lässt sich konstatieren, dass ältere Probanden mit einem höheren Jobalter in der vorliegenden Stichprobe eine geringere Bereitschaft zur Weiterempfehlung der Produkte/Dienstleistungen aufwiesen als jüngere Befragte bzw. Berufsanfänger. Gleiches gilt auch für die Studierendenstichprobe in Bezug auf ihre Studiendauer. So konnten vor allem unter Studienan-

fängern Promoters aufgefunden werden. Die Zusammenhänge zwischen der Weiterempfehlungsbereitschaft und diesen soziodemographischen Aspekten fielen aber, auch wenn sie signifikant wurden, eher gering aus ($r < -.25$). Überraschenderweise wiesen die weiteren explorativ untersuchten soziodemographischen Variablen keine Verbindung zur Loyalität auf. Gleiches galt auch für die Persönlichkeit der Befragten.

Neben diesen generellen Befunden eigneten sich die weiteren erfassten unabhängigen Variablen in unterschiedlichem Maße als korrelative Treiber in Abhängigkeit der jeweiligen NPS-Operationalisierung.

So erwiesen sich die Variablen Arbeitszufriedenheit ($r = .79, p < .001$), Mitarbeiterorientierung ($r = .76, p < .001$), evaluatives Unternehmensimage ($r = .76, p < .001$) sowie Vertrauen zur vollständigen Organisation ($r = .76, p < .001$) als besonders starke Treiber in Bezug auf die *Bereitschaft, das Unternehmen als Arbeitgeber weiterzuempfehlen*. Die durchgeführte hierarchische Regressionsanalyse unterstrich die Bedeutung dieser Aspekte, wobei allein die Konstrukte Zufriedenheit, Betriebsklima und Image 73% der Varianz der Loyalität aufklären konnten.

Die in Bezug auf die *Produkte/Dienstleistungen* berechneten Zusammenhänge kristallisierten das evaluative Image ($r = .65, p < .001$), die globale Zufriedenheit ($r = .64, p < .001$), das globale Klima ($r = .62, p < .001$) sowie das Vertrauen zur Organisation ($r = .62, p < .001$) als *bedeutsamste Korrelate* heraus. Dieser Befund konnte durch die durchgeführte Regressionsanalyse bestätigt werden, die eine Varianzaufklärung von 54% aufzeigen konnte.

In der *Studierendenstichprobe* erwiesen sich die Aspekte evaluatives Image ($r = .73, p < .001$), affektives Commitment ($r = .72, p < .001$), Studienendenorientierung ($r = .71, p < .001$) und wahrgenommene Dynamik ($r = .70, p < .001$) als bedeutendste Loyalitätstreiber. So konnte auch allein anhand dieser Konstrukte 65% der Varianz der abhängigen Variablen erklärt werden.

Treiber, die eine Differenzierung der Loyalitätsgruppen ermöglichen

Eine *Differenzierung aller Loyalitätsgruppen* erlaubten die Konstrukte Betriebsklima, affektives Commitment, Vertrauen und Image, während sich die Passives und Promoters nur partiell bezüglich der Konstrukte Zufriedenheit und Fluktuationsintention unterschieden. Eine untergeordnete Rolle spielten in diesem Zusammenhang auch die normative Verpflichtung der Befragten sowie ihre kalkulative Bindung.

Wie im Falle der korrelativen Analysen unterschieden sich die untersuchten unabhängigen Variablen auch bezüglich ihrer Differenzierungskraft in Abhängigkeit der NPS-Variante. So führten die berechneten Diskriminanzanalysen zu einer korrekten Klassifikationsrate von 74% bezüglich der Bereitschaft, das Unternehmen als Arbeitgeber weiterzuempfehlen. Eine Hitrate

von 77% in Bezug auf die Bereitschaft, das Unternehmen als Produzent/Dienstleister weiterzuempfehlen, belegte die Eignung der potentiellen Treiber im Unternehmensbereich, während in der Studierendenstichprobe sogar 80% der Befragten den von Reichheld (2003) konzipierten Gruppen korrekt zugewiesen werden konnten.

Die Ergebnisse der zwei verfolgten Fragestellungen geben erste Hinweise darauf, dass eine allgemeine Erhöhung der Weiterempfehlungsbereitschaft von einer differenzierten Beeinflussung der Auftretenswahrscheinlichkeit von Promoters bzw. Detractors unterschieden werden muss. Dieser Befund konnte ebenfalls durch die qualitative Tier-Assoziation bestätigt werden. Dabei zeigte eine Analyse der Antworten, dass allein die Detractors Tiere mit ihrem Unternehmen assoziiert haben, die von keiner anderen Gruppe genannt wurden. Dabei führten die Detractors zu einem Tiere auf, die phobische Reize darstellen wie Schlange und Spinne sowie zu anderen Tiere, die gesellschaftlich negativ besetzt sind (Esel, Kamel, Faultier). Darüber hinaus assoziierten die Illoyalen solche Tiere, die mit Gefahr oder Tod assoziiert werden, wie beispielsweise Geier, Hyäne, Krokodil und Wolf.

Ferner konnten verschiedene Angebote identifiziert werden, deren Vorhandensein die Weiterempfehlungsbereitschaft positiv zu beeinflussen scheint. Die Mitarbeiter, deren Unternehmen die beruflichen Angebote Qualitätszirkel, Dienstreisen, Weiterbildungsmaßnahmen und Tagungen/Vorträge anboten, empfahlen ihr Unternehmen mit einer signifikant größeren Wahrscheinlichkeit sowohl als Arbeitgeber als auch als Produzent/Dienstleister weiter als die Befragten, deren Unternehmen nicht über derartige Angebote verfügten. Ähnliche Befunde konnten auch bei der Studierendenstichprobe festgestellt werden. So empfahlen die Studenten ihre Hochschulen eher weiter, wenn diese Stammtische, Fach- und Berufsberatungen sowie Exkursionen anboten.

Diskussion und Handlungsimplicationen

Der vorliegende Beitrag liefert mit der durchgeführten Treiberanalyse erste Ansätze für denkbare Handlungsimplicationen, da einige mögliche Loyalitätsdeterminanten auf statistischer Ebene identifiziert werden konnten. Über alle drei NPS-Operationalisierungen hinweg stand die Loyalität insbesondere mit dem *Image*, dem *Vertrauen* und dem *Klima* der jeweiligen Institution in Zusammenhang. Die fehlende Abhängigkeit der Weiterempfehlungsbereitschaft von persönlichen Variablen verdeutlicht die große potenzielle Beeinflussungsmöglichkeit der Loyalität seitens der Unternehmen/Hochschulen.

Da die Besonderheit des von Reichheld (2003) entwickelten Loyalitätsmaßes – im Vergleich zu gängigen Zufriedenheitsmaßen – in einer Fokussierung der Personen mit einer positiven/negativen Weiterempfehlungsbereit-

schaft besteht, während die Neutralen/passiv Zufriedenen vernachlässigt werden, lässt sich vermuten, dass auch im Bereich interner Kunden das Hauptaugenmerk auf eine gezielte indirekte Beeinflussung der Gruppenzugehörigkeit gelegt werden sollte. Tabelle 2 stellt die Eignung der verschiedenen untersuchten unabhängigen Variablen als Loyalitätstreiber (in Abhängigkeit der zwei verfolgten Fragestellungen) resümierend dar.

Tabelle 2: Überblick über die Treiberqualität der untersuchten unabhängigen Variablen

		NPS-A		NPS-P/D		NPS-ST	
		korrelativ	diff.	korrelativ	diff.	korrelativ	diff.
<i>Soziodemographie</i>	(Job)Alter bzw. Studiendauer	0	Pa > P –	Pa > P –		D > P	
<i>Persönlichkeit</i>	individuelle Unterschiede	0	/	0	/	/	0
<i>Zufriedenheit</i>	global	++		++		+	
	Ansehen	+		+		+	
	Einkommen/ Studiengebühren	+		+		+	
	Kollegiales Verhältnis	+		+		+	P > D
	Entwicklungsmöglichkeiten	+		+		+	
	Selbstverwirklichung	+		+		+	
	Selbstständigkeit	+		+		+	P > D
	Arbeitsaspekte/ Studieninhalte	+		+		+	
	Arbeitsplatzsicherheit/Attraktivität HS-Standort	+		+		+	

		NPS-A		NPS-P/D		NPS-ST	
		korrelativ	diff.	korrelativ	diff.	korrelativ	diff.
<i>Betriebs- klima</i>	global	++		++		++	
	Mitarbeiter/ Studentenorientierung	++		++		++	
	Unternehmens-/ Hochschulpolitik	+		++		++	
	Information und Mitsprache	+		+		+	
	Karriereförderung	+		+		/	/
	Arbeitszeit	+		+		/	/
	Arbeitsinhalte bzw. Lehre	+		+		+	
	Kundenorientierung	+		+		/	/
	Kollegen bzw. Kommilitonen	+		+		+	
	Vorgesetzte bzw. Dozenten	++		+		+	
	Entlohnung	+		+		/	/

		NPS-A		NPS-P/D		NPS-ST	
		korrelativ	diff.	korrelativ	diff.	korrelativ	diff.
<i>Commitment</i>	affektiv	++		++		++	
	normativ	+		+		+	
	kalkulativ	0		0		0	
<i>Vertrauen</i>	zur Organisation/ Hochschule	++		++		++	
	zum Vorgesetzten/ Dozenten	++		+		+	
<i>Bindung</i>	Bleibeintention	++		+		+	
<i>Image</i>	global	++		++		++	
	Evaluation	++		++		++	
	Aktivität	+		+		++	
	Potenz	+		+		+	
<i>Angebote</i>	Freizeitangebote	0		0		0	
	Berufliche Angebote	0		0		0	

Legende:

	Detractors < Passives < Promoters	Diff- = differenzierende Treiber	++	starker Zusammenhang
	Detractors < Passives = Promoters	korrelativ = korrelative Treiber	+	moderater Zusammenhang
	Detractors = Passives = Promoters	P = Promoters Pa = Passives D = Detractors	-	negativer Zusammenhang
	Vorhandensein > nicht-Vorhandensein		0	kein Zusammenhang

Wie dieser Überblick verdeutlicht, sollten Unternehmen/Hochschulen, die den NPS ihrer internen Kunden positiv verändern möchten, insbesondere am *Betriebs-/Hochschulklima* ansetzen. Dieses Konstrukt verfügt über eine hohe Differenzierungskraft in Bezug auf die Detractors, Passives und Promoters und dürfte eine gezielte institutionsspezifische Ableitung von Handlungsimplicationen ermöglichen. Gemäß den vorliegenden Befunden ist dabei vor allem die wahrgenommene Orientierung an den Mitarbeitern/Studenten von Bedeutung sowie eine transparente Politik, die die Werte der Institutionsangehörigen berücksichtigt und überzeugende Leitbilder betont. Ein persönli-

cher Umgang unter den Kollegen/Kommilitonen sowie ein als gerecht und unterstützend erlebtes Verhalten der Vorgesetzten/Dozenten sollten ferner zu einem positiven Klima beitragen. Auch die Güte des *Vertrauenskonstruktes* als Loyalitätstreiber betont die Wichtigkeit einer von Kooperation geprägten Atmosphäre, die sich durch gegenseitige Investitionen auszeichnet. In diesem Zusammenhang hat sich in der vorliegenden Arbeit insbesondere die Existenz von *Angeboten* als fruchtbar erwiesen, welche die berufliche Entwicklung interner Kunden sowie ein arbeits-/studiumsbezogenes Zusammengehörigkeitsgefühl fördern. Es kristallisierte sich ferner heraus, dass finanzielle Anreize die Weiterempfehlungsbereitschaft nicht bedeutend beeinflussen. Loyalität ist somit *nicht käuflich*, sondern sollte vielmehr durch immaterielle Anreize gewonnen werden, die zur Identifikation mit der Institution einladen. Diesbezüglich hat sich auch das *Image* als bedeutender Loyalitätstreiber herausgestellt, wobei weniger die wahrgenommene Potenz von Bedeutung scheint als das Bild eines aktiven und sozialen Unternehmens. Schließlich basiert eine hohe Weiterempfehlungsbereitschaft auf einer *affektiven Verbundenheit*, die nicht durch hohe Wechselbarrieren erzwungen werden kann.

Literatur

- BAILLOD, J. & Semmer, N. (1994). Fluktuation und Berufsverläufe bei Computerfachleuten. *Zeitschrift für Arbeits- und Organisationspsychologie*, 38 (4), 152-163.
- CHAMBERS, E., Foulon, M., Handfield-Jones, H., Hankin, S. & Michaels, E. (1998). The war for talent. *McKinsey Quarterly*, 3, 44 -57.
- FELFE, J. & Six, B. (2006). Die Relation von Arbeitszufriedenheit und Commitment. In L. Fischer (Hrsg.), *Arbeitszufriedenheit. Konzepte und empirische Befunde* (S. 37-60). Göttingen: Hogrefe.
- FELFE, J., Six, B. & Schmook, R. (2002). Fragebogen zur Erfassung von affektivem, kalkulatorischem und normativem Commitment gegenüber der Organisation, dem Beruf / der Tätigkeit und der Beschäftigungsform (COBB). In A. Glöckner-Rist (Hrsg.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente*. (Version 7.00). Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- GANESH, J., Arnold, M. & Reynolds, K. (2000). Understanding the customer base of service providers: An examination of the differences between switchers and stayers. *Journal of Marketing*, 64, 65-87.
- GEBERT, D. & von Rosenstiel, L. (2002). *Organisationspsychologie* (5. Aufl.). Stuttgart: Kohlhammer.
- GIERING, A. (2000). *Der Zusammenhang zwischen Kundenzufriedenheit und Kundenloyalität: eine Untersuchung moderierender Effekte*. Wiesbaden: Gabler.
- GOMEZ, P. (2000). Loyalität...ein Abschied? *Index*, 1, 52-53.
- GRAEFF, P. (1998). *Vertrauen zum Vorgesetzten und zum Unternehmen*. Berlin: Wissenschaftlicher Verlag.

- HAASE, D. (1997). *Organisationsstruktur und Mitarbeiterbindung. Eine empirische Analyse in Kreditinstituten*. Köln: Dt. Inst.-Verl.
- IAB Kurzbericht. (2006). *Aktuelle Analysen aus dem Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit*, (Ausgabe Nr. 6). Zugriff am 10.09.2009 unter <http://doku.iab.de/kurzber/2006/kb2206.pdf>
- JUDGE, T., Heller, D. & Mount, M. (2002). Five-factor model of personality and job satisfaction: a meta-analysis. *Journal of Applied Psychology*, 87, 237-249.
- LANGER, M., Ziegele, F. & Hennig-Thurau, T. (2001). *Hochschulbindung – Entwicklung eines theoretischen Modells, empirische Überprüfung und Ableitung von Handlungsempfehlungen für die Hochschulpraxis*. Gütersloh: Centrum für Hochschulentwicklung. Zugriff am 10.01. 2007 von www.che.de/downloads/hs_bindung_bericht0105.pdf
- LENECKE, K. (2005). *Hochschulbindung durch Student Services. Grundlagen, Analysen, Perspektiven*. Saarbrücken: Dr. Müller.
- NGUYEN, N. & LeBlanc, G. (2001). Image and reputation of higher education institutions in students' retention decisions. *International Journal of Educational Management*, 15 (6), 303-311.
- ORGAN, D. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.
- OSGOOD, C., Suci, G. & Tannenbaum, P. (1975). *The measurement of meaning*. Urbana, Ill: Univ. Press.
- RAMMSTEDT, B. & John, P. (2006). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1), 203-212.
- REICHELIED, F. (2003). The one number you need to grow. *Harvard Business Review*, December, Product no. 5534.
- REICHELIED, F. (2006). *The ultimate question. Driving good profits and true growth*. Boston: McGraw-Hill.
- RINDERMANN, H. & Oubaid, V. (1999). Auswahl von Studienanfängern durch Universitäten – Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 20 (3), 172-191.
- SCHERRER, A. (1975). *Das Phänomen der Mund-zu-Mund-Werbung und seine Bedeutung für das Konsumentenverhalten*. Unveröffentlichte Dissertation. Universität Freiburg (Schweiz).
- SCHWAIGER, M. (2003). Der Student als Kunde – eine empirische Analyse der Zufriedenheit Münchner BWL-Studenten mit ihrem Studium. *Beiträge zur Hochschulforschung*, 25 (1), 32-62.
- STUMPF, H., Angleitner, A., Wieck, Th., Jackson, D. & Beloch-Till, H. (1985). *Deutsche Personality Research Form (PRF)*. Göttingen: Hogrefe.
- SUNDARAM, D., Mitra, K. & Webster, C. (1998). Word-of-mouth communications: A motivational analysis. *Advances in Consumer Research*, 25 (1), 527-531.
- TINTO, V. (1975) Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45, 89-127.

VON Rosenstiel, L. (1992). Führungs- und Führungsnachwuchskräfte: Spannungen und Wandlungen in Phasen gesellschaftlichen Umbruchs. *Zeitschrift für Personalforschung*, 6 (3), 327-351.

WITHEY, M. & Cooper, W. (1992). What's loyalty? *Employee Responsibilities and Rights Journal*, 5 (3), 231-240.

Teil II
Evaluation und Qualitätssicherung in
verschiedenen Anwendungsfeldern

Die Reihe der sieben Beiträge zu diesem Themenbereich wird mit zwei Artikeln aus dem breiten Spektrum der ZEM-Aktivitäten zur Evaluation kultureller und kulturpolitischer Programme (für diverse in der Einleitung benannte außeruniversitäre Auftraggeber) eröffnet, nämlich zu *Entsendungsprozessen*¹ und zur *Museumsevaluation*. Die weiteren Artikel stammen aus Projekten (gefördert von der EU, der Bundesanstalt für Straßenwesen, der DFG) zur Evaluation von Maßnahmen und zur Entwicklung von entsprechenden Messinstrumenten in den – bei manchen dieser Projekte miteinander verknüpften – Bereichen *Verkehr/Mobilität, Altern und Gesundheit*.

1 *Kursive Passagen* benennen bzw. umschreiben die Themen der Artikel

Kristina Heidemann

Perspektiven zur Evaluation von Entsendungsprozessen und Auslandseinsätzen – die Entwicklung eines längsschnittlichen Designs

Zusammenfassung

Der Prozess einer Auslandsentsendung ist eine komplexe Angelegenheit, die für alle Beteiligten (entsendende Organisation, Mitarbeiter, Kollegen, Familie etc.) in organisatorischer und persönlicher Hinsicht mit vielfältigen Herausforderungen verbunden ist. Bei der Evaluierung solcher Maßnahmen bleibt häufig unklar, unter welchen Bedingungen der Auslandsaufenthalt für wen ein Erfolg war. Eine am ZEM etablierte Forschungsgruppe »Expatriates« hat sich zum Ziel gesetzt, die Anforderungen verschiedener Personengruppen zu definieren, zu integrieren und eigene Evaluationskonzepte zu entwickeln. Dabei wurde ein längsschnittlicher Ansatz entwickelt: Geprüft wird, welche Maßnahmen im Rahmen des Entsendungsprozesses wann zur Lebens- und Arbeitszufriedenheit, dem Commitment und der Einstellung zum Gastland des Entsendeten beitragen und welche Lerneffekte bei den Entsandten auftreten. Die hier beschriebenen Ergebnisse stellen die gekürzte Version einer Dissertation dar. Die ausführliche Version ist im Buchhandel (Dr. Müller Verlag) verfügbar. Die übrigen Ergebnisse der Arbeitsgruppe sind bei der PACE erschienen und unter www.zem.uni-bonn.de aufgeführt.

Einleitung

Globalisierung ist eines der zentralen Schlagworte, das die sozialen und wirtschaftlichen Transformationsprozesse der westlichen Industriestaaten kennzeichnet (Clement, 2002). Bedingt durch die Zunahme der internationalen Zusammenarbeit wird der Aufenthalt von Personen im Ausland verstärkt erforderlich (u.a. Bergemann, 1994; Brandenburger, 1995; Deresky, 1994). Bolino und Feldmann (2000) berichten, dass zunehmend Personal von Konzernen ins Ausland geht und die Mitarbeiter¹ dort einige Monate oder Jahre in Niederlassungen oder anderen Partnerorganisationen arbeiten; Yan, Zhu und Hall (2002) schätzen die Zahl solcher Mitarbeiter in den USA auf ca.

1 Im Folgenden wird, der besseren Lesbarkeit halber, auf die Erwähnung männlicher und weiblicher Formen verzichtet. Gleichmaßen sind – wenn nicht ausdrücklich anderes gekennzeichnet – beide Formen gemeint.

150.000. In der Entwicklungszusammenarbeit zeigt sich ein ähnliches Bild: Wenngleich zunehmend nationale Kräfte eingebunden werden, spielt der Aufenthalt westlicher Fachkräfte im Empfängerland eine große Rolle für die Gestaltung der Projekte (Bundesministerium für Entwicklungszusammenarbeit, BMZ, 2004). In Abgrenzung zu Flüchtlingen, Auswanderern und anderen Personen, die aus verschiedenen Gründen ihren Heimatstaat verlassen, bezeichnen Dowling und Schuler (1990) Personen, die sich aus *beruflichen* Gründen für einen begrenzten Zeitraum im Ausland aufhalten, als *Expatriates*: »an expatriate is a professional/managerial employee moved from a country to, and for employment in, another country.« (S.171).

Theoretischer Hintergrund

Die Entsendung von Expatriates findet bei Profit- oder Non-Profit-Organisationen statt. Als Auslandseinsatz oder -aufenthalt wird ein zeitlich begrenzter aber längerfristiger Aufenthalt jenseits der eigenen Staatsgrenzen bezeichnet (Geistmann, 2002). Kennzeichen eines solchen Auslandsaufenthaltes ist, dass innerhalb einer bestimmten Zeit mehr oder weniger fest umrissene Aufgaben bearbeitet werden sollen. In der Wirtschaft gelten die Vermittlung von Fach- und Führungswissen und die Einrichtung von geeigneten Koordinations- und Kommunikationsprozessen als Hauptziele von Auslandsentsendungen, ebenso die Vertretung von Interessen und die Einrichtung von Steuerungs- und Controllingprozessen (Welge & Holtbrügge, 2001). Auch die Aufgaben der sogenannten Fachkräfte für Entwicklungszusammenarbeit bestehen in der Vermittlung von Wissen, der Verbesserung der Kommunikation verschiedener Interessensgruppen und in der Evaluation bzw. dem Controlling der Vorhaben, indem die Wirksamkeit von Projekten und das Finanzkonzept überprüft werden (BMZ, 2004). Schließlich wird der Auslandseinsatz in der Wirtschaft auch dazu eingesetzt, ein Reservoir an auslandserfahrenen Mitarbeitern zu entwickeln und die Kompetenz des Einzelnen, mit Mitgliedern anderer Kulturen umzugehen, zu stärken (Müller, 1996; Üstünsöz-Beurer, 1998). Auch der Deutsche Entwicklungsdienst (DED) wirbt für das Engagement als Entwicklungshelfer mit der Erweiterung des Kompetenzprofils während des Auslandsaufenthaltes und einer gezielten Steigerung von Belastbarkeit, Flexibilität und Kommunikationsfähigkeit (DED, 2007). Kammel und Paul (1993) berichten, dass zudem die Entwicklung der Persönlichkeit bei Expatriates häufig ein Hauptmotiv für die Entsendung ist.

Die Entsendung von Personal ist ein vielschichtiger und komplexer Vorgang, dessen Bestandteile sich in vielerlei Hinsicht gliedern lassen. Zur besseren Orientierung innerhalb des Entsendungsprozesses werden in Bezug auf *Zeit*, *beteiligte Personen* und *Inhalte* verschiedene Ebenen unterschieden, die in Tabelle 1 dargestellt sind. Aufgrund der Individualität der Akteure sind der

Vielfalt der Problembereiche keine Grenzen gesetzt, so dass immer nur eine Auswahl betrachtet werden kann.

Tabelle 1: Gliederung des Entsendungsprozesses in Phasen, beteiligte Akteure und Problembereiche

<i>Phasen</i>	<i>Akteure</i>	<i>Problembereiche</i>
<i>Auswahl & Entscheidung</i>	<i>Expatriate</i>	<i>interkulturelle</i>
<i>Vorbereitung</i>	<i>Partner/-in</i>	<i>berufliche</i>
<i>Ausreise & Betreuung</i>	<i>Kinder</i>	<i>administrative</i>
<i>Rückkehr & Reintegration</i>	<i>Entsendende Organisation</i>	<i>rechtliche</i>
	<i>Aufnehmende Organisation</i>	<i>finanzielle</i>
	<i>Kollegen im Heimat- & Gastland</i>	<i>mit der Alltagsgestaltung verbundene</i>
	<i>Umfeld</i>	<i>sprachliche</i>
		<i>persönliche</i>
		<i>etc...</i>

Der Ansatz des ZEM stellt den Expatriate in den Mittelpunkt der Betrachtungen, versucht dabei aber dessen Schwierigkeiten mit den verschiedenen Akteuren und in den unterschiedlichen inhaltlichen Bereichen zu berücksichtigen. Die verschiedenen Zeitabschnitte werden in besonderer Weise berücksichtigt, da das entwickelte Untersuchungsdesign den gesamten Entsendungsprozess abbildet. So wird ermöglicht, dass Expatriates in den unterschiedlichen Phasen ihrer Entsendung zu ihrer aktuellen Lage befragt und rückblickende oder kurzfristige Bewertungen von Problembereichen vermieden werden.

Problematisch für die Evaluation von Auslandsaufenthalten ist, dass nicht immer festgestellt werden kann, ob das Ziel der Entsendung erreicht wurde und inwieweit dies in Zusammenhang mit den Leistungen des Expatriates zu bringen ist, insbesondere bei Fehlschlägen und frühzeitigen Abbrüchen: So kann es geschehen, dass ein Expatriate seine Aufgabe gut erfüllt, eine geplante Kooperation aber aus Gründen, die an anderen Stellen im Unternehmen entschieden werden, dennoch nicht zustande kommt. Gleichmaßen kann eine Kooperation zwar hervorragend verlaufen, die Arbeit des Expatriates leistet aber keinen nennenswerten Beitrag zu diesem Verlauf. Verlässt ein guter Mitarbeiter, der seine Aufgabe erfolgreich bewältigt hat, nach der Rückkehr das Unternehmen und nimmt dabei sein gesammeltes Know-how mit, kann dies langfristig auch nicht als Erfolg gewertet werden (Hirsch, 1996; Winter, 1996). Ebenso kann die Entsendung dazu beitragen, die persönliche Situation eines Expatriates erheblich zu verschlechtern (Groß, 1994). Um die Wahrscheinlichkeit eines langfristig erfolgreichen Verlaufs der Entsendung zu erhöhen, werden vielfältige Maßnahmen zur Auswahl, Vorbereitung, Begleitung und Rückkehr ergriffen; es liegen aber verschiedene Ansichten und Befunde darüber vor, wann welche Maßnahmen eingesetzt

werden sollen, um den Auslandsaufenthalt zu einem Erfolg werden zu lassen. Wie wichtig es aber ist, die Anforderungen an den Aufenthalt und die ihn begleitenden Maßnahmen genau zu definieren, verdeutlichen die Aussagen von Blake und Heslin (1983) am Beispiel interkulturellen Trainings: Der Beitrag eines interkulturellen Trainings zum Gelingen des Auslandsaufenthaltes kann nur dann identifiziert werden, wenn klar definiert ist, was im Ausland erreicht werden soll und inwieweit das interkulturelle Training dazu beiträgt, diese Ziele besser zu erreichen. Gleichermaßen können nur Aussagen über die Eignung und Bedeutsamkeit der gesamten Entsendungsgestaltung getroffen werden, wenn zuvor festgelegt wird, was genau im Ausland überhaupt erreicht werden soll, also soweit möglich eine Wirkungskette aufgelegt wird.

Fragestellung

Im Folgenden werden zunächst die einzelnen Maßnahmen im Rahmen von Auswahl, Vorbereitung, Betreuung und Reintegration zum Evaluationsgegenstand gemacht und mit dem Ziel verbunden, die Entsendung aus Sicht der Expatriates möglichst optimal zu gestalten. Der Zweck der Evaluation besteht darin, Aussagen darüber zu ermöglichen, welche der vielen Maßnahmen sich besonders eignen, den Auslandseinsatz zum Erfolg zu machen, und entsprechende Handlungsalternativen zu empfehlen. Die unterschiedliche Bewertung von Maßnahmen in Abhängigkeit vom Zeitpunkt sowie der Position (Hieronymi, 2001; Winter, 1996) macht deutlich, dass eine Zeitperspektive bei der Bewertung der Maßnahmen erforderlich ist und Zufriedenheitsmessungen, z.B. unmittelbar nach einem interkulturellen Training, nicht unbedingt Rückschlüsse über die Bedeutsamkeit der Aussage für den Erfolg des Auslandsaufenthaltes zulassen.

Geprüft wurden deshalb verschiedene Fragen:

- Wie hilfreich werden die verschiedenen Maßnahmen, die von den entsendenden Organisationen im Verlauf der Zeit angeboten wurden, wahrgenommen? Bei Expatriates, denen die jeweilige Maßnahme nicht angeboten wurde, wird zu den verschiedenen Zeitpunkten der Entsendung erfragt, ob sie das Angebot *vermisst haben*. So kann deutlich werden, wie die verschiedenen Maßnahmen zu den unterschiedlichen Zeitpunkten der Entsendung bewertet werden.
- Wie hängen die Maßnahmen der Auswahl, Vorbereitung, Betreuung und Rückkehr neben anderen Gestaltungsmerkmalen der Auslandsentsendung mit der Ausprägung der Lebens- und Arbeitszufriedenheit, dem Commitment sowie der Einstellung zum Gastland, hier die Operationalisierung des erfolgreichen Auslandsaufenthaltes, zusammen?

Anschließend wird die Entwicklung von Persönlichkeit und Kompetenzen der Expatriates gesondert betrachtet und geprüft, inwieweit diese stattfindet. Evaluationsgegenstand ist hier der Auslandsaufenthalt insgesamt, dessen häufig formuliertes Ziel darin besteht, das Potenzial der Expatriates zu erweitern (u.a. Bittner & Reisch, 1996). Zweck der Evaluation in diesem Zusammenhang ist die verbesserte Nutzung der im Ausland erworbenen Fähigkeiten und Fertigkeiten.

Untersuchungsdesign

Der Fragestellung liegt ein komplexes Design zugrunde, welches den Auslandsaufenthalt in drei Phasen unterteilt und so im Querschnittsdesign die Abbildung des gesamten Entsendungsprozesses ermöglicht: Befragt wurden

- Expatriates zu Beginn der Entsendung (Gruppe 1)
- während der Entsendung (Gruppe 2) und
- bei der Rückkehr ins Heimatland (Gruppe 3).

Alle Gruppen wurden zu *zwei* Erhebungszeitpunkten an der Untersuchung beteiligt, die ca. sechs Monate auseinander lagen (Messwiederholung). Die Teilnehmer in Gruppe 1 wurden vor ihrer Ausreise (Gruppe 1.1) und unmittelbar nach Beginn des Auslandsaufenthaltes befragt (Gruppe 1.2), Gruppe 2 hielt sich zu beiden Erhebungszeitpunkten im Gastland auf (Gruppen 2.1 und 2.2). Die Teilnehmer aus Gruppe 3 standen zum ersten Erhebungszeitpunkt kurz vor dem Abschluss ihres Auslandsaufenthaltes (Gruppe 3.1) und befanden sich zum zweiten Erhebungszeitpunkt wieder in ihrem Heimatland (Gruppe 3.2).

Für die detaillierte Evaluation der den Auslandsaufenthalt begleitenden Maßnahmen stehen demnach Daten aus sechs Erhebungszeitpunkten über den gesamten Verlauf des Auslandsaufenthaltes zur Verfügung (unabhängige Betrachtung der Erhebungen ohne Berücksichtigung der Messwiederholung). Die Evaluation des Auslandseinsatzes insgesamt in Bezug auf Veränderungen von Kompetenzen, Soft Skills und Persönlichkeitseigenschaften in allen Phasen der Entsendung beruht auf Daten aus zwei Erhebungen an der gleichen Stichprobe (drei Gruppen mit jeweils zwei Messzeitpunkten).

Methode

Für den hier beschriebenen Untersuchungsansatz wurden verschiedene Herangehensweisen gewählt. Bei der Beurteilung der Maßnahmen stehen die Befragungszeitpunkte, differenziert nach Maßnahmen der Profit- und Non-Profit-Gruppe, unverbunden nebeneinander, wie in Tabelle 2 dargestellt.

Tabelle 2: Unverbundene Messzeitpunkte

$X_{1,1}$	$X_{1,2}$	$X_{2,1}$	$X_{2,2}$	$X_{3,1}$	$X_{3,2}$
→					
Zeit					
Gruppe 1.1	Gruppe 1.2	Gruppe 2.1	Gruppe 2.2	Gruppe 3.1	Gruppe 3.2
Vor der Entsendung (Heimatland)	Beginn der Entsendung (Ausland)	Mitte der Entsendung (beide Gruppen Ausland)	Ende der Entsendung (Ausland)		Nach der Entsendung (Heimatland)

Das Ziel dieser Betrachtungen besteht lediglich darin, einen Überblick über die wahrgenommene Bedeutung der verschiedenen Maßnahmen im Verlauf der Entsendung zu erhalten, um so die Substanz von Beurteilungen spezifischer Maßnahmen erhöhen zu können, wenn sich geeignete Beurteilungszeitpunkte ermitteln lassen.

Anschließend wurde geprüft, wie die Vielzahl der im Zusammenhang mit Auslandsentsendungen diskutierten Maßnahmen, Eigenschaften, Soft Skills und Kompetenzen mit der Ausprägung der abhängigen Variablen *Lebens- und Arbeitszufriedenheit*, *Commitment* und *Einstellung zum Gastland* in Zusammenhang stehen. Ziel der Überprüfung ist wiederum, Aussagen über die Bedeutsamkeit einzelner Maßnahmen zu den verschiedenen Zeitpunkten des Entsendungsablaufs zu ermöglichen. Auch für diese Fragestellungen werden die Messzeitpunkte unverbunden nebeneinander gestellt.

Statistisch betrachtet wird an dieser Stelle geprüft, ob die Ausprägung der Untersuchungsteilnehmer in Bezug auf die unabhängigen Variablen die Stichprobe in zwei unterschiedliche Populationen teilt. Formuliert wird am Beispiel der Begleitung des Expatriates durch seine Familie folgende statistische Hypothese:

$$\begin{aligned} &\mu_{\text{mit Familie Zeitpunkt 1, Gruppe 1}} > \mu_{\text{ohne Familie Zeitpunkt 1, Gruppe 1}} \cap \mu_{\text{mit Familie Zeitpunkt 2,}} \\ &\text{Gruppe 1} > \mu_{\text{ohne Familie Zeitpunkt 2, Gruppe 1}} \cap \mu_{\text{mit Familie Zeitpunkt 1, Gruppe 2}} > \mu_{\text{ohne Fa-}} \\ &\text{milie Zeitpunkt 1, Gruppe 2}} \cap \mu_{\text{mit Familie Zeitpunkt 2, Gruppe 2}} > \mu_{\text{ohne Familie Zeitpunkt 2, Gruppe 2}} \\ &\cap \mu_{\text{mit Familie Zeitpunkt 1, Gruppe 3}} > \mu_{\text{ohne Familie Zeitpunkt 1, Gruppe 3}} \cap \mu_{\text{mit Familie Zeit-}} \\ &\text{punkt 2, Gruppe 3}} > \mu_{\text{ohne Familie Zeitpunkt 2, Gruppe 3}} \end{aligned}$$

Nicht alle Fragestellungen in diesem Zusammenhang lassen sich jedoch mit Mittelwertunterschieden hinreichend prüfen, so dass in einigen Fällen auch Korrelationen berechnet werden. Statistisch gesehen lautet die Hypothese, im Beispiel formuliert für den Zusammenhang der Lebenszufriedenheit mit der Lernbereitschaft:

$$\rho_{\text{Lebenszufriedenheit Zeitpunkt 1, Gruppe 1 Lernbereitschaft Zeitpunkt 1, Gruppe 1}} > 0 \cap \rho_{\text{Lebenszufriedenheit Zeitpunkt 2, Gruppe 1 Lernbereitschaft Zeitpunkt 2, Gruppe 1}} \cap \rho_{\text{Lebenszufriedenheit Zeitpunkt 1, Gruppe 2 Lernbereitschaft Zeitpunkt 1, Gruppe 2}} \cap \rho_{\text{Lebenszufriedenheit Zeitpunkt 2, Gruppe 2 Lernbereitschaft Zeitpunkt 2, Gruppe 2}} \cap \rho_{\text{Lebenszufriedenheit Zeitpunkt 1, Gruppe 3 Lernbereitschaft Zeitpunkt 1, Gruppe 3}} \cap \rho_{\text{Lebenszufriedenheit Zeitpunkt 2, Gruppe 1 Lernbereitschaft Zeitpunkt 2, Gruppe 1}}^2$$

Im Anschluss wurde der Auslandsaufenthalt als Ganzes in den Fokus der Aufmerksamkeit gerückt. Die zuvor als *unabhängige Variablen* eingeführten Konstrukte werden im zweiten Hypothesenteil als *abhängige Variablen* betrachtet, die durch die Ausreise ins Gastland beeinflusst werden. Dabei wird geprüft, ob, bedingt durch die Teilnahme am Auslandsaufenthalt, Veränderungen in den Ausprägungen der Kompetenzen, Soft Skills und Eigenschaften auftreten, wie von entsendender Organisation und Expatriate häufig erwartet wird und theoretisch aus vielerlei Gründen abzuleiten ist. Dabei werden Mittelwertunterschiede *innerhalb* der verschiedenen Untersuchungsgruppen *zwischen* den Untersuchungszeitpunkten geprüft und postuliert, dass sich die zu den beiden Untersuchungszeitpunkten erhobenen Mittelwerte voneinander unterscheiden (Taris, 2000). Verwendet werden können in diesem Abschnitt lediglich solche Datensätze, bei denen die Untersuchungsteilnehmer beide Fragebögen in ihrer Gruppe beantwortet haben, sofern keine systematische Selektion – beispielsweise besonders unzufriedener Untersuchungsteilnehmer – stattgefunden hat. Statistisch ergibt sich folgende Hypothese: $\mu_{\text{Zeitpunkt 1}} \neq \mu_{\text{Zeitpunkt 2}}$

Neben den anhand bestehender Fragebögen ermittelten Eigenschaften, Kompetenzen und Soft Skills (*indirekte* Selbsteinschätzungen) wurden die Untersuchungsteilnehmer gebeten, sich zusätzlich durch *direkte* Selbsteinschätzungen zu bewerten. Dabei wurde neben der Selbsteinschätzung zum *aktuellen* Zeitpunkt auch eine *prospektive* Einschätzung der Ausprägung bestimmter Eigenschaften, Kompetenzen und Soft Skills erbeten, jeweils mit einem konkreten Zeitpunkt in der Vergangenheit bzw. Zukunft verbunden (vgl. Tabelle 3), um die Erwartung an eine potenzielle Veränderung zu erfassen.

2 Neben der Querschnittskorrelation zwischen den zu betrachtenden Variablen wurde eine Cross Lagged Panel Correlation (CLPC) berechnet, deren Ergebnisse durch die Angabe von kausalen Prioritäten (Rudinger & Bierhoff, 1980) oder kausalen Richtungen (Erdfelder & Undorf, 2003) Hinweise darauf liefern, welche der beiden Merkmalsausprägungen die dominierende ist, also die inhaltliche Richtung der Korrelation angibt (Erdfelder, Rietz & Rudinger, 1996). Diese lieferte keine relevanten Ergebnisse.

Tabelle 3: Schematische Darstellung der Erhebungszeitpunkte der Selbsteinschätzungen

Entsendungsverlauf	Messzeitpunkt	Entscheidungsphase	Beginn der Entsendung	Ende der Entsendung	Rückkehrphase
Gruppe 1 Entsendungskandidaten	1. Befragungszeitpunkt	A	P		
	2. Befragungszeitpunkt		A		
Gruppe 2 Expatriates vor Ort	1. Befragungszeitpunkt		A	P	
	2. Befragungszeitpunkt			A	
Gruppe 3 Rückkehrer	1. Befragungszeitpunkt			A	P
	2. Befragungszeitpunkt				A

A = Selbsteinschätzung zum Befragungszeitpunkt; P = prospektive Einschätzung

Untersuchungsmaterial

Für die vorliegende Fragestellung wurden insgesamt zwölf verschiedene Fragebögen entwickelt, die in weiten Teilen übereinstimmend formuliert, aber auf die Besonderheiten der jeweiligen Gruppe in Bezug auf Profit- bzw. Non-Profit-Bereich und Entsendungszeitpunkt abgestimmt waren. Alle Fragebögen gliedern sich in fünf Abschnitte und wurden den Teilnehmern online zur Verfügung gestellt. Unterschieden werden Fragen zur *Arbeit im Gastland* (Commitment, Arbeitszufriedenheit, Beziehung zur entsendenden Organisation, globale, gesellschaftliche und organisationale Kompetenz), zur *Kultur* (Sprache, interkulturelle Kompetenz, Einstellung gegenüber dem Gastland sowie der Übereinstimmung zwischen den eigenen Werten und denen des Gastlandes, Sicherheit in der Kommunikation mit Angehörigen des Gastlandes) und zum *Erleben der eigenen Person* (Belastbarkeit, Dominanz, Flexibilität, Lernbereitschaft, Engagement, Organisationstalent, Ehrgeiz, Entscheidungssicherheit, emotionale Stabilität, Lebenszufriedenheit, Gesundheit, familiäre und soziale Unterstützung). Vorangehend werden jedoch einige Fragen zum *Entsendungsverlauf* gestellt und im ersten Abschnitt die *soziodemografischen Variablen und die Rahmenbedingungen der Entsendung* erhoben. Die Grundlage der eingesetzten Erhebungsmaterialien bilden in der Regel stark gekürzte Versionen bereits bestehender, mehrfach geprüfter Fragebögen verschiedener Autoren zu den jeweiligen Kompetenzen, Persönlichkeitseigenschaften und Soft Skills. Dabei wurden sowohl einzelne *Skalen* als auch *Items* aus den bestehenden Fragebögen entnommen (Rost & Hoberg,

1996; Franke, 1997). Die Fragebögen wurden einem Pretest unterzogen und die Skalenhomogenität mit befriedigenden Ergebnissen geprüft.

Stichprobe

An der Befragung haben insgesamt 375 Personen aus 22 verschiedenen Organisationen des Profit- und Non-Profit-Bereichs teilgenommen. 68,3% lassen sich dem Non-Profit-Bereich zuordnen, 31,7% dem Profit-Bereich. Die an der Untersuchung beteiligten Expatriates waren im Durchschnitt 34,7 Jahre alt (Standardabweichung: 8,7). Insgesamt waren 39,8% der Untersuchungsteilnehmer weiblich, 60,2% männlich. Im Profit-Bereich liegt der Frauenanteil bei 7,0% ($n = 7$), was, geht man davon aus, dass Frauen sich nicht systematisch weniger an der Untersuchung beteiligt haben, widerspiegelt, dass Organisationen im Profit-Bereich in der Regel Männer ins Ausland entsenden, wie schon in der Literatur beschrieben (Groß, 1994; Waldhoff, Hieronymi & Rietz, 2003). Insgesamt verfügen 81,2% der Teilnehmer über die (Fach-)Hochschulreife. Die meisten befragten Expatriates haben zudem ein Studium abgeschlossen. Die Teilnehmer hielten sich in insgesamt 81 verschiedenen Ländern auf. Meist beherrschten sie die Landessprache und verfügten über längere Auslandserfahrung.

Die Verteilung der Ergebnisse lässt erkennen, dass im Profit-Bereich vorwiegend Kandidaten ins Ausland entsendet werden, die in einer festen Partnerschaft leben. Im Non-Profit-Bereich ist die Verteilung weitgehend ausgeglichen, ebenso viele Teilnehmer leben nicht in fester Partnerschaft. Viele Teilnehmer haben Kinder, deutlich wird aber, dass vor allem im Profit-Bereich meist Männer mit Kindern entsendet werden. Ebenso wird aufgezeigt, dass Frauen nur selten von ihren Kindern und Partnern begleitet werden; die mit ausreisenden Partner waren in der Regel weiblich.

Ergebnisse³ und Diskussion

Zunächst kann festgehalten werden, dass die Stichprobe, die der Untersuchung zugrunde liegt, in wesentlichen Merkmalen mit denen übereinstimmt, die in anderen Studien beschrieben werden, so dass die Ergebnisse sich gut in die bisherigen Befunde einreihen können. Hinweise auf systematische Verzerrungen oder Drop-outs konnten nicht gefunden werden. Als Besonderheit kann gelten, dass sowohl Teilnehmer aus dem Profit- als auch dem Non-Profit-Bereich sowie Teilnehmer in vielen Ländern und Kontinenten beteiligt wurden.

3 Auf die Darstellung ausgewählter, detaillierter Ergebnisse wird hier aus Platzgründen verzichtet und wiederum auf die ausführliche Darstellung in der Dissertation und ihren Anhang verwiesen.

Wahrnehmung der Maßnahmen

Zur Analyse der subjektiven Wahrnehmung verschiedener Maßnahmen zur Vorbereitung und Begleitung von Aufenthalt und Rückkehr lagen keine Hypothesen vor. Ziel der Betrachtung war es, herauszufinden, ob die einzelnen Maßnahmen der entsendenden Organisation durchgängig positiv oder negativ beurteilt werden oder ob sich zu den verschiedenen Entsendungszeitpunkten deskriptiv unterschiedliche Ausprägungen ergeben. Mittelwertvergleiche wurden nicht durchgeführt, da die verschiedenen Teilnehmergruppen sich zu heterogen zusammensetzen und insofern Unterschiede in den Ausprägungen nicht immer auf den unterschiedlichen Zeitpunkt im Entsendungsprozess zurückzuführen sind, sondern möglicherweise auch auf unterschiedliche Maßnahmen. Festgehalten werden muss zunächst, dass die meisten Teilnehmer die von der entsendenden Organisation ergriffenen Maßnahmen in allen Phasen der Entsendung als recht hilfreich empfanden. Die Urteile derer, denen die unterschiedlichen Maßnahmen nicht angeboten wurden, fielen zudem in Bezug auf die meisten Maßnahmen eindeutig aus: Die Untersuchungsteilnehmer waren sich recht einig, dass die entsprechenden Maßnahmen gefehlt haben.

Aus den Ergebnissen konnte aber abgeleitet werden, dass die Berücksichtigung verschiedener Zeitpunkte bei der Evaluation entsprechender Maßnahmen durchaus ihre Berechtigung hat: Konrad, Hertel und Behr (2002) beschreiben, dass die Evaluation interkultureller Trainings beispielsweise in der Regel unmittelbar nach der Durchführung durch die Erfassung der Zufriedenheit mit der Maßnahme erfolgt, und auch Delfs, Heidemann und Steckhan (2007) bestätigen die Beschränkung auf die Feststellung der Zufriedenheit mit der Tätigkeit der entsendenden Organisation während und nach dem Auslandsaufenthalt als »Evaluationsmaßnahme« in Unternehmen verschiedener Größen. Vor allem für die Vorbereitungsmaßnahmen kann dieses Vorgehen aufgrund der ermittelten Ergebnisse nicht als hilfreich angesehen werden, da viele Teilnehmer sich zu diesem Befragungszeitpunkt eines Urteils über die Unterstützung durch die Vorbereitungsmaßnahmen enthalten. Dies ist nachvollziehbar, fällt es doch auch augenscheinlich schwer, anzugeben, wie hilfreich ein interkulturelles Training war, wenn man noch keine Gelegenheit hatte, das Training mit der Realität abzugleichen, stellt aber dennoch keinen trivialen Befund dar, da eben in der Regel entsprechende Fragen nicht zu einem späteren Zeitpunkt gestellt werden. Will man also die Effektivität eines interkulturellen Trainings erfassen, sollte dies möglichst nach Beginn des Auslandsaufenthaltes erfolgen sowie erneut nach einer längeren Eingewöhnungszeit.

Die Maßnahmen zur Begleitung der Rückkehr bereits im Rahmen der Vorbereitung wurden mit Fortschreiten des Auslandsaufenthaltes (auf deskriptiver Ebene) positiver bewertet. Bezüglich der Unterstützung bei

beruflichen oder privaten Schwierigkeiten wird im (deskriptiven) Vergleich der Messzeitpunkte deutlich, dass diejenigen, die nicht unterstützt wurden, dies zunehmend einforderten – vermutlich in Abhängigkeit zur im Verlauf der Zeit wachsenden Anzahl derjenigen, die Schwierigkeiten erlebt und in der konkreten Situation Unterstützung vermisst haben. Hier kann den entsendenden Organisationen mehr Aufmerksamkeit für die Schwierigkeiten ihrer Angehörigen nahe gelegt werden. Zudem kann wiederum geschlussfolgert werden, dass die Evaluation solcher Maßnahmen dann erfolgen sollte, wenn ihre Bedeutung für den Expatriate offensichtlich wird.

Einige Maßnahmen der entsendenden Organisation wurden über den gesamten Entsendungszeitraum konstant als (eher) hilfreich bezeichnet, zum Beispiel die Vorbereitung durch Kontakte zu zurückgekehrten Kollegen oder anderen Personen, die sich bereits im Gastland aufgehalten haben, die Vorbereitung der Angehörigen oder das Informationsmaterial. Hier ist möglicherweise insgesamt die *Qualität* der Maßnahmen ausschlaggebend – Informationen und vermittelte Kontakte wurden entweder als hilfreich erlebt oder eben nicht. Deutlich wurde im Zusammenhang mit den Kontakten, dass sie vermisst wurden, wenn sie nicht vermittelt wurden. Ihnen wird demzufolge von den Teilnehmern eine hohe Unterstützungsfunktion zugeschrieben.

Von allen Teilnehmern wurden mehr Kontakte zu Angehörigen der eigenen Firma gewünscht. Dies bestätigt die Befunde von Hieronymi (2001), nach denen Expatriates am liebsten durch einen Trainer auf den Auslandsaufenthalt vorbereitet werden, der das Gastland und die entsendende Organisation kennt. Bezüglich der *Bedeutsamkeit* der Vorbereitung und Betreuung insgesamt bleibt die Frage offen, inwieweit diese gewünscht und sinnvoll ist: Wurden Maßnahmen ergriffen, wurden sie in den meisten Fällen als hilfreich oder eher hilfreich bezeichnet, unterblieben Maßnahmen, wurden sie aber nicht durchgehend vermisst. Vermutlich sind also die in den Organisationen häufig bereits getroffenen Regelungen, bei Bedarf individuell im privaten Bereich einzugreifen (Delfs, Heidemann & Steckhan, 2007), im Sinne der Teilnehmer. Ist eine Unterstützung im privaten Bereich wie bei den Profit-Organisationen einmal etabliert, sollten die Leistungen auch stringent angeboten werden.

Beitrag einzelner Maßnahmen zum Entsendungserfolg

Als wichtigstes Ergebnis kann festgehalten werden, dass sich viele der in der internationalen Literatur diskutierten Merkmale des Entsendungsprozesses (z.B. Beteiligung der Familie, Beziehungen zu Gastlandangehörigen, Beherrschen der Sprache, Ausprägung verschiedener Persönlichkeitseigenschaften) als nicht relevant für die hier definierten Indikatoren des Entsendungserfolgs erwiesen und damit insgesamt nahe legen, die Maßnahmen rund um den Auslandsentsendungsprozess zu überdenken. In der hier beschriebe-

nen Studie erwies sich keine der im Zuge der einzelnen Maßnahmen getroffenen Entscheidungen über alle Gruppen hinweg konstant als bedeutsam für die Ausprägung der Erfolgsindikatoren.

Nichtsdestotrotz lassen sich aus der Arbeit viele Befunde für die Gestaltung von Auslandsentsendungen ableiten.

In Bezug auf die *Auswahl* der Expatriates konnte festgestellt werden, dass bei der Auswahl von Expatriates der Schwerpunkt der Anforderungen im Bereich der Persönlichkeitseigenschaften auf die emotionale Stabilität und die Belastbarkeit gelegt werden kann, sowie auf die Sicherheit im Umgang mit den Gastlandangehörigen, da diese Variablen die stabilsten Zusammenhänge zu den Erfolgsfaktoren aufwiesen. Entgegen der lauter werdenden Forderung, auch die Familie in die Entsendungsplanung einzubeziehen (Groß, 1994; Puttlitz et al., 2003; Tung, 1981, 1984; Waldhoff, Hieronymi & Rietz, 2003) und vor allem Mitarbeiter mit Familie zu entsenden, erwies sich hier die familiäre Situation im Gastland nicht als bedeutsam für die Ausprägung der Lebenszufriedenheit: Es konnten keine Unterschiede zwischen denen festgestellt werden, die durch ihren Partner oder ihre Kinder ins Ausland begleitet wurden und denen, die alleine ausreisten oder nicht in einer festen Partnerschaft lebten.

Auch die Auslandserfahrung hatte keine systematischen Effekte auf die Ausprägung der Lebenszufriedenheit oder die Einstellung zum Gastland, gleichermaßen konnte kein Zusammenhang zwischen dem Beherrschen der Gastlandssprache und der Lebenszufriedenheit oder der Einstellung zum Gastland festgestellt werden. Inwieweit es aber sinnvoll wäre, diese Faktoren bei einem Auswahlgespräch nicht zu berücksichtigen, ist mindestens diskussionswürdig.

Im Zusammenhang mit der *Vorbereitung* des Auslandsaufenthaltes zählt zu den wichtigsten Ergebnissen, dass sich die Maßnahmen der länderspezifischen Vorbereitung (Sprachkurs, interkulturelles Training und Look&See Trip) im Profit- und im Non-Profit-Bereich negativ auf die Ausprägung der Lebenszufriedenheit, der Arbeitszufriedenheit, des Commitments und der Einstellung zum Gastland auswirkten: Untersuchungsteilnehmer, die an den verschiedenen Maßnahmen teilgenommen haben, wiesen eine geringere Ausprägung auf den jeweiligen Konstrukten bzw. eine negativere Einstellung zum Gastland auf als ihre Kollegen, die entsprechende Maßnahmen nicht mitgemacht haben. Die Effekte traten zwar nicht in allen Gruppen auf, sind aber so konstant in allen geprüften Zusammenhängen vertreten, dass sie nicht ignoriert werden können, zumal die beobachteten Effekte auf den ersten Blick insbesondere für die Einstellung zum Gastland der Absicht der Maßnahmen vollständig entgegen wirken und der kulturellen Vorbereitung auf das Gastland in der Literatur große Bedeutung beigemessen wird. Glanz (2003) ebenso wie Vance und Paik (2002) vertreten allerdings die Ansicht, dass die meist auf der Grundlage von Kulturschockparadigmen entwickelten

interkulturellen Trainings der Realität nicht entsprechen und Expatriates nicht auf die Komplexität ihrer Aufgaben vorbereiten. Vor allem die nur innerhalb eines Tages vermittelten Informationen seien nicht hilfreich; stattdessen eigene sich vor allem die tägliche Unterstützung der Expatriates bei Schwierigkeiten, verbunden mit Erklärungen zu interkulturell bedingten Ursachen dieser Schwierigkeiten.

Für die Vorbereitung durch Kontakte zu Kollegen oder Expatriates aus anderen Organisationen ließen sich keine signifikanten Effekte auf die Lebenszufriedenheit oder die Arbeitszufriedenheit feststellen. Das Commitment erwies sich als höher, wenn die Teilnehmer in einzelnen Gruppen als hilfreich erlebte Kontakte zu ihren Kollegen hatten; die Commitmentwerte der Teilnehmer, die solche Kontakte vermissten, waren entsprechend geringer ausgeprägt. Dies spricht dafür, dass Kontakte vor allem dann durch die entsendende Organisation vermittelt werden sollten, wenn die Kollegen bereit und in der Lage sind, den zukünftigen Expatriates als hilfreich empfundene Tipps zu geben. Einschränkend ist darauf hinzuweisen, dass solche vereinzelt auftretenden Signifikanzen nicht überbewertet werden sollten, zumal die subjektive Qualität der Vorbereitung durch Kontakte weder bei der Lebenszufriedenheit noch bei der Arbeitszufriedenheit zu signifikanten Ergebnissen führte.

Bei den Maßnahmen der beruflichen Vorbereitung hat vor allem der Mentor bzw. dessen Fehlen Einfluss auf die Ausprägung der Lebens- und Arbeitszufriedenheit sowie des Commitments. Untersuchungsteilnehmer, die einen Mentor vermissten, wiesen im Profit- und Non-Profit-Bereich während des Auslandsaufenthaltes eine geringere Lebenszufriedenheit auf als ihre Kollegen, die einen Mentor als nicht notwendig erachteten, darüber hinaus gegen Ende des Aufenthaltes und im Gesamtwert eine geringere Arbeitszufriedenheit und im Gesamtwert ein geringeres Commitment. Zudem scheint es ausschlaggebend zu sein, dass ein Mentor als hilfreich wahrgenommen wird; seine bloße Existenz reicht nicht aus, Effekte zu evozieren.

Für die *Betreuung* zeigten sich keine stabilen, signifikanten Unterschiede in der Ausprägung der Erfolgsindikatoren aufgrund der Teilnahme bzw. Nicht-Teilnahme an einzelnen Maßnahmen. Werden alle Ergebnisse zur Betreuung insgesamt betrachtet, wird deutlich, dass die Ausprägungen der Arbeitszufriedenheit und des Commitments sensibler auf die berufliche Unterstützung reagieren als die Lebenszufriedenheit, was sich durchaus sinnvoll interpretieren lässt und für die grundsätzliche Bedeutung der Maßnahmen der beruflichen und auch privaten Betreuung spricht. Es scheint, dass das anhaltende Interesse der entsendenden Organisation am Mitarbeiter, das sich in der Betreuung der Expatriates auch während ihres Aufenthaltes manifestiert, sich positiv auf das Commitment und die Arbeitszufriedenheit auswirkt. Darüber hinaus zeigte sich, dass die fehlende Unterstützung bei der Eingewöhnung ins Gastland zu Beginn des Auslandsaufenthaltes sowohl auf die Lebenszufrie-

denheit als auch auf das Commitment einen mindernden Einfluss ausübt – auch hier kann die entsendende Organisation solche Einbrüche vermeiden.

Als besonders relevant für die Ausprägung der Arbeitszufriedenheit erwies sich in diesem Zusammenhang die Gestaltung der Arbeitsaufgabe: Klare Kompetenzrichtlinien und klare Angaben zur Arbeitsaufgabe korrelierten im Non-Profit-Bereich nahezu durchgängig signifikant mit der Arbeitszufriedenheit (nur nach der Rückkehr zeigten sich keine bedeutsamen Ergebnisse); das gleiche Muster zeigte sich zur Klarheit der Arbeitsaufgabe im Profit-Bereich. Auch die Korrelationen mit der Lebenszufriedenheit und dem Commitment stützen insgesamt die Bedeutung solcher Vorgaben durch die entsendende Organisation. In Anbetracht des von Hieronymi et al. (2003) sowie Wirth (1992) berichteten Bedürfnisses der Expatriates nach einer herausfordernden Aufgabe, bei der man sich außerhalb der Vorgaben der Organisation bewegen kann, verwundert diese hohe Bedeutung klar formulierter Aufgaben und Richtlinien. Möglicherweise müssen sich die entsendende Organisation und der Expatriate auf ein Mindestmaß an Vorgaben einigen, um beide Seiten zufrieden zu stellen. Solche Vorgaben würden auch die von Schneider et al. (2003) beschriebenen Defizite bei der Nutzung des im Ausland erworbenen Wissens erleichtern, da so bereits im Vorhinein deutlicher definiert würde, welche Aufgaben der Expatriate im Einzelnen übernimmt und welche Kompetenzen damit verbunden sind.

Möglicherweise kann die Abweichung der Ergebnisse von den bisherigen Studien auf die Auswahl der gewählten Erfolgsindikatoren (hohe Ausprägung von Arbeits- und Lebenszufriedenheit, Commitment sowie eine positive Einstellung zum Gastland) zurückgeführt werden, da für die ausgewählten Variablen keine hohe Übereinstimmung über den gesamten Entsendungszeitraum darstellbar war oder die Operationalisierung der Variablen auslandsspezifischer hätte erfolgen müssen. Durch die Ermittlung von Mittelwertsunterschieden werden zudem möglicherweise signifikante Zusammenhänge auf individueller Ebene verdeckt. Auch Besonderheiten des Untersuchungsdesigns und der heterogenen Zusammenstellung der Untersuchungsstichprobe können Effekte verdecken.

Beitrag des Auslandsaufenthaltes zur Entwicklung von Kompetenzen, Soft Skills und Persönlichkeitseigenschaften

Ein zweiter Schwerpunkt der vorliegenden Untersuchung liegt auf der Feststellung, inwieweit der Aufenthalt (und letztlich alle in seinem Verlauf ergriffenen Maßnahmen, geknüpften Kontakte usw.) dazu beitragen kann, die von Expatriates und entsendender Organisation erwarteten Veränderungen von Soft Skills, Persönlichkeitseigenschaften und Kompetenzen anzustoßen. Eine Studie von Delfs, Heidemann und Steckhan (2007) verdeutlicht, dass die meisten entsendenden Organisationen keine genauen Vorstellungen über

die Art der Entwicklung ihrer Mitarbeiter haben. Auch Schneider et al. (2003) berichten von nur wenig differenzierten Erwartungen der von ihnen befragten Unternehmen. Die Arbeit außerhalb des eigenen Landes wird meist allgemein mit der Entwicklung von Fähigkeiten verbunden (NFTC et al., 2002), ohne genau zu definieren, *was* konkret dazu beiträgt, *welche* Veränderungen zu induzieren (Glanz, 2003). Angenommen wurde hier, dass sich positive Entwicklungen verschiedener Eigenschaften und Kompetenzen mit Fortschreiten des Auslandsaufenthaltes manifestieren würden. Eine solche Entwicklung konnte aber nur in Einzelfällen nachgewiesen werden, insgesamt konnten keine messbaren Verbesserungen festgestellt werden.

Ergänzend zur indirekten Erfassung der Merkmale durch Fragebögen wurden in der vorliegenden Studie zu beiden Untersuchungszeitpunkten in jeder Gruppe direkte Selbsteinschätzungen erhoben: Die Teilnehmer wurden gebeten, ihre Ausprägung auf den oben beschriebenen Soft Skills, Kompetenzen und Persönlichkeitseigenschaften zum *momentanen* Zeitpunkt auf einer Skala von 1 = *sehr gering ausgeprägt* bis 6 = *sehr stark ausgeprägt* einzuschätzen. Zusätzlich wurden die Teilnehmer zum ersten Erhebungszeitpunkt gebeten, *prospektiv* einzuschätzen, wie sie sich nach Beginn der Entsendung (Gruppe 1), im weiteren Verlauf des Auslandsaufenthaltes (Gruppe 2) bzw. nach der Rückkehr ins Heimatland (Gruppe 3) selbst einschätzen würden. Die so gewonnenen Daten wurden verwendet, um zu prüfen, ob Untersuchungsteilnehmer Veränderungen *erwarteten* (Vergleich zwischen momentanen und prospektiven Selbsteinschätzungen) und ob sie selbst solche Veränderungen auch an sich *beobachten* konnten (Vergleich der momentanen Selbsteinschätzungen vom ersten zum zweiten Erhebungszeitpunkt). Die Ergebnisse haben gezeigt, dass die meisten Teilnehmer in allen Gruppen positive Veränderungen ihrer Kompetenzen erwarteten, vor allem im Zusammenhang mit der interkulturellen Kompetenz, diese aber kaum an sich beobachteten, so dass die vorliegenden Ergebnisse dafür sprechen, die betrachteten Merkmale als stabil zu betrachten.

Bei der Auslegung der Ergebnisse kann angeführt werden, dass eventuell die ausgewählten Zeitpunkte zu kurzfristig sind, um Veränderungen in den unterschiedlichen Variablen real ausbilden zu können – in den Gruppen 1 und 3 sind allerdings die kritischen Ergebnisse Ein- bzw. Ausreise eingebunden, die als Auslöser für Veränderungen diskutiert werden (Filipp, 1990; Mummendey, 1995). Geeigneter wäre möglicherweise der Vergleich mit einem Zeitpunkt weiter vor Beginn der Entsendung, da vielleicht die gedankliche Auseinandersetzung mit dem Auslandsaufenthalt schon Veränderungen bei den Untersuchungsteilnehmern auslöst. Gleichmaßen kann eine Erhebung mit größerem Abstand zum Ende des Auslandsaufenthaltes besser geeignet sein, wenn die Rückkehrer mit größerem Abstand zum Entsendungsprozess mehr Zeit hatten, sich selbst in ihrer alten Umgebung zu erleben und Veränderungen möglicherweise besser realisieren können.

Eine andere Möglichkeit zur Erklärung der fehlenden Effekte kann in der Auswahl der betrachteten Variablen liegen – denkbar ist, dass Veränderungen sich auf anderen Variablen abbilden. Gegebenenfalls wäre auch eine Fokussierung auf konkretere Fähigkeiten notwendig; allerdings müssten dann auch andere Erhebungsmethoden gewählt werden, da kaum Instrumente zur Selbsteinschätzung fachlicher Fähigkeiten zur Verfügung stehen.

Ergänzend wäre hilfreich, auch Urteile von Kollegen, Vorgesetzten oder Partnern einzuholen, um ein umfassenderes Bild zur Persönlichkeit der Expatriates zu bekommen. Als Forderung lässt sich analog zu Schneider et al. (2003) festhalten, dass die entsendenden Organisationen ihre Erwartungen an den Auslandsaufenthalt ebenso wie die Mitarbeiter klar formulieren, um so Anhaltspunkte dafür zu liefern, welche Eigenschaften und Kompetenzen sich im Ausland überhaupt entwickeln sollen. Dies würde die Durchführung von solchen Untersuchungen erleichtern und würde für Expatriates und ihre begleitenden Partner deutlicher machen, welchen persönlichen und fachlichen Gewinn sie durch einen Auslandsaufenthalt erzielen können.

Fazit

Abschließend kann festgehalten werden, dass die vorliegende Untersuchung entsendenden Organisationen vielfältige Möglichkeiten aufgezeigt hat, vor allem das Commitment, aber auch die Arbeits- und Lebenszufriedenheit sowie die Einstellung ihrer Expatriates zum Gastland positiv zu beeinflussen. Dies geschieht vor allem darüber, dass die entsendende Organisation den Expatriates zu allen Zeitpunkten des Entsendungsverlaufs ihr Interesse an dessen Arbeit signalisiert, bei Schwierigkeiten zur Verfügung steht und darüber hinaus den Einsatz und v.a. die Aufgabe sorgfältig plant.

Für die Veränderung der beschriebenen Kompetenzen, Soft Skills und Persönlichkeitseigenschaften wird eine erneute Untersuchung mit einer homogeneren Gruppe, einem echten Längsschnitt und größeren Abständen zur Aus- bzw. Einreise vorgeschlagen – versehen mit dem Hinweis, dass eine solche Stichprobe nur sehr schwer zu rekrutieren ist. Sinnvoll scheint auch, konkretere Fähigkeiten und Fertigkeiten der Expatriates zu erfassen. Die hier ermittelten Ergebnisse geben keinen Anlass, von einer Entwicklung der Eigenschaften der Expatriates auszugehen; zudem konnten weder für die interkulturelle Kompetenz, den Umgang mit den Gastlandangehörigen, noch die globale, organisationale oder gesellschaftliche Kompetenz positive Auswirkungen des Auslandsaufenthaltes festgestellt werden.

Grundsätzlich kann für die Durchführung weiterer Evaluationen im Zusammenhang mit verschiedenen Aspekten der Auslandsentsendung – gerade durch die Abweichung der Ergebnisse von den bisher diskutierten – ein hoher Bedarf abgeleitet werden. Auch die Betrachtung der Maßnahmen über den gesamten Entsendungsverlauf hat gezeigt, dass Evaluationen im Verlauf des

Entsendungsprozesses deutlichere Anhaltspunkte zu ihrer Wirksamkeit und Gestaltung geben können und sich insofern zur Ableitung von Verbesserungsvorschlägen anbieten. Zunächst ist es definitiv notwendig, möglicherweise über die Arbeits- und Lebenszufriedenheit, das Commitment und die Einstellung zum Gastland hinaus, festzulegen, was aus Sicht der an einem Auslandsaufenthalt beteiligten Akteure als Entsendungserfolg gilt. Die vorliegende Untersuchung hat mit der Berücksichtigung aller Phasen des Entsendungsprozesses vielleicht neue Anstöße dazu geliefert.

Literatur

- BERGEMANN, N. (1994). Interkulturelles Management. In A. Thomas (Hrsg.), *Psychologie und multikulturelle Gesellschaft* (S. 247-250). Göttingen: Verlag für Angewandte Psychologie.
- BITTNER, A. & Reisch, B. (1996). *Anforderungen an Auslandsmanager: Erfolgs- und Misserfolgskriterien*. Bad Honnef: Institut für interkulturelles Management.
- BLAKE, B.F. & Heslin, R. (1983). Evaluating cross-cultural training. In D. Landis & R.W. Brislin (Eds.), *Handbook of intercultural training*, (Vol. 1, pp.203-233). New York: Pergamon.
- BOLINO, M. & Feldman, D. (2000). The antecedents and consequences of underemployment among expatriates. *Journal of Organizational Behavior*, 21 (8), 889 – 911.
- BRANDENBURGER, M. (1995). *Interkulturelles Management. Ein Konzept zur Entsendung von Führungskräften unter besonderer Berücksichtigung von Auswahl und Vorbereitung*. Köln: Botermann & Botermann.
- BUNDESMINISTERIUM für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ). (2004). *Medienhandbuch 2002*. Zugriff am 01.07.2004 unter <http://www.bmz.de/infothek/fachinformationen/allgefachinfo/Medienhandbuch2002.pdf>
- CLEMENT, U. (2002). Flexibilisierung als Zielbegriff und Zauberwort in Wirtschaft und Ausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 98 (3), 383-404.
- DEUTSCHER Entwicklungsdienst (DED). (2007). *Als Fachkraft im Ausland*. Zugriff am 26.02.2007 unter http://www.ded.de/cipp/ded/custom/pub/content,lang,1/oid,72/ticket,g_u_e_s_t/~Fachkraft_im_Ausland.html
- DELFS, M.M., Heidemann, K. & Steckhan, H. (2007). Evaluation von Auslandsentsendungen: Wie ermitteln Unternehmen nach der Rückkehr eines Expatriates, ob der Auslandseinsatz erfolgreich war? In K. Heidemann, H. Steckhan & C. Rietz (Hrsg.), *Erfolg von Auslandseinsätzen: Neue Beurteilungsansätze. Einfluss auf die Persönlichkeit und Kriterien zur Erfolgsbeurteilung* (Schriftenreihe der Arbeitsgruppe »Expatriates« am Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn, S. 80-96). Bonn: PACE.

- DERESKY, H. (1994). *International management: Managing across borders and cultures*. HarperCollins: New York.
- DOWLING, P.J. & Schuler, R.S. (1990). *International dimensions of human resource management*. Boston, MA.: PWS-Kent.
- FILIPP, S.H. (Hrsg.). (1990). *Kritische Lebensereignisse* (2. Aufl.). München: Psychologie Verlags Union.
- FRANKE, G. H. (1997). *Effekte der computerunterstützten Anwendung bei klinisch-psychodiagnostischen Selbstbeurteilungsinventaren unter besonderer Berücksichtigung der Itemreihenfolge. Experimentelle Studien*. Essen: Universitätsklinikum.
- GEISTMANN, C. (2002). *Interkulturelle Kompetenz – eine wichtige und förderbare Fähigkeit in der internationalen Zusammenarbeit. Entwicklung eines Konzeptes zur schrittweisen Förderung interkultureller Kompetenz aufgezeigt am Beispiel der Handelspartner Frankreich und Deutschland*. Norderstedt: Books on Demand GmbH.
- GLANZ, L. (2003). Expatriate stories: A vehicle of professional development abroad? *Journal of Managerial Psychology*, 18 (3), 259-274.
- GROB, P. (1994). *Die Integration der Familie beim Auslandseinsatz von Führungskräften – Möglichkeiten und Grenzen international tätiger Unternehmen* (Dissertation an der Hochschule St. Gallen). Hallstadt: Rosch.
- HIERONYMI, K. (2001). *Die Ex-Ante-Evaluation der Einsatzvorbereitung von Expatriates*. Unveröffentlichte Diplomarbeit, Rheinische Friedrich-Wilhelms-Universität Bonn.
- HIERONYMI, K., Rietz, C., Lissner, K. & Melin, N. (2003). *Die Ermittlung von Kriterien zur Erfassung des Erfolges bei Auslandsentsendungen* (Schriftenreihe der Arbeitsgruppe »Expatriates« am Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn). Bonn: PACE.
- HIRSCH, K. (1996). Reintegration von Auslandsmitarbeitern. In N. Bergemann & A.L.J. Sourisseaux (Hrsg.), *Interkulturelles Management* (2. Aufl., S. 201-228). Heidelberg: Physica-Verlag.
- KAMMEL, A. & Paul, S. (1993). 'Re-Entry': Aufgabe des Personalmanagements multinationaler Unternehmen. *Personal*, 10, 450 – 454.
- KONRAD, U., Hertel, G. & Behr, B. (2002). Interkulturelle Management-Trainings: Eine Bestandsaufnahme von Konzepten, Methoden und Modalitäten in Deutschland. *Zeitschrift für Sozialpsychologie*, 33 (4), 197-207.
- MÜLLER, S. (1996). Auslandsorientierung als Zielsetzung der Personalentwicklung. In A. Thomas (Hrsg.), *Psychologie interkulturellen Handelns* (S. 341-364). Göttingen: Hogrefe.
- MUMMENDEY, H.-D. (1995). Selbstkonzept-Änderungen nach kritischen Lebensereignissen. In S.-H. Filipp (Hrsg.), *Kritische Lebensereignisse* (3. Aufl., S. 252-269). Weinheim: Beltz, Psychologie Verlags Union.

- NFTC, SHRM and Cendant Mobility. (2002). *New approaches to global mobility: 2002 worldwide benchmark study*. Zugriff am 13.06.2006 unter: <http://www.cendantmobility.com>
- PUTTLITZ, N., Lissner, K., Heidemann, K. & Rietz, C. (2003). *Auslandsentsendung von Expatriates – Wenn die Familie zurückbleibt. Eine Pilotstudie über die Situation betroffener Familien* (Schriftenreihe der Arbeitsgruppe »Expatriates« am Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn). Bonn: PACE.
- ROST, D.H. & Hoberg, K. (1996). *Itempositionsveränderungen in Persönlichkeitsfragebogen: Methodischer Kunstfehler oder tolerierbare Praxis?* Marburg: Philipps-Universität.
- SCHNEIDER, N.F., Bennwitz, J., Karsch, A. & Oersterling, C. (2003). Auslandsmobilität optimal managen. *Personalwirtschaft*, 5, 40-44.
- TARIS, T.W. (2000). *A primer in longitudinal data analysis*. London: Sage Publications.
- TUNG, R. (1981). Selection and training of personnel for overseas assignment. *Columbia Journal of World Business*, 16, 68-78.
- TUNG, R.L. (1984). Strategic management of human resources in the multinational enterprise. *Human Resource Management*, 23, 129-143.
- ÜSTÜNSÖZ-Beurer, D. (1998). *Auslandspraktika als Instrumente einer europäischen Berufsqualifizierung*. Dissertation, FernUniversität-Gesamthochschule Hagen.
- VANCE, C.M. & Paik, Y. (2002). One size fits all in expatriate pre-departure training? Comparing the host country voices of Mexican, Indonesian and US workers. *Journal of Management Development*, 21 (7), 557-71.
- WALDHOFF, A., Hieronymi, K. & Rietz, C. (2003). *Die Bewertung des Entsendungsprozesses von Expatriates durch den mitausreisenden Partner* (Schriftenreihe der Arbeitsgruppe »Expatriates« am Zentrum für Evaluation und Methoden an der Rheinischen Friedrich-Wilhelms-Universität in Bonn). Bonn: PACE.
- WELGE, M.K. & Holtbrügge, D. (2001). *Internationales Management*. Landsberg/Lech: Verlag Moderne Industrie.
- WINTER, G. (1996). Reintegrationsproblematik: Vom Heimkehren in die Fremde und vom Wiedererlernen des Vertrauten. In A. Thomas (Hrsg.), *Psychologie interkulturellen Handelns* (S. 365-381). Göttingen: Hogrefe.
- WIRTH, E. (1992). *Mitarbeiter im Auslanseinsatz: Planung und Gestaltung*. Wiesbaden: Gabler.
- YAN, A., Zhu, G. & Hall, D.T. (2002). International assignment for career building: A model of agency relationships and psychological contracts. *Academy of Management Review*, 27, 373-391.

Maximilian Ponert & Uwe Kleinemas

Wissenschaftliche Museumsevaluation: Ein Bericht erfolgreicher Kooperationen zwischen dem Zentrum für Evaluation und Methoden und Museen der Bonner Region¹

Im Unterschied zum nordamerikanischen Sprachraum stellen Evaluationen von Museumsausstellungen in Europa und in Deutschland immer noch eher eine Ausnahme dar. Während in den USA bereits seit den 70er Jahren Evaluationskonzepte in diesem Bereich entwickelt wurden, werden im deutschsprachigen Raum die Museumsbesucher nur wenig durch Evaluationsprojekte in den Planungs- und Gestaltungsprozess einer Ausstellung mit einbezogen (AfeB e.V., 2003). Mögliche Gründe hierfür sind häufig Zweifel an der Verwertbarkeit der Ergebnisse, Furcht vor negativen Untersuchungsergebnissen oder schlicht ein generelles Misstrauen in das Instrument Evaluation (Friedman, 1996). Auch vermeintlich hohe Kosten können ein Grund sein, eine Ausstellung nicht evaluativ zu begleiten. Hierzu stellt Schäfer (2004) allerdings heraus, dass die Evaluation einer Ausstellung nur durchschnittlich einen Anteil von 1-2% des Gesamtbudgets ausmacht.

Vor diesem Hintergrund ist die Bedeutung der nun bereits mehrjährigen und erfolgreichen Kooperation des Zentrums für Evaluation und Methoden (ZEM) der Rheinischen Friedrich-Wilhelms-Universität mit sowohl dem Rheinischen Landesmuseum (RLM) als auch der Kunst- und Ausstellungshalle der Bundesrepublik Deutschland (KAH) in Bonn im Bereich Kulturevaluation besonders hoch einzuschätzen. In der KAH wurden unter anderem die monografisch-wissenschaftsgeschichtliche Ausstellung »Alexander von Humboldt – Netzwerke des Wissens« (1999/2000), die archäologisch-kulturhistorische Ausstellung »Gold aus dem alten Peru – Die Königsgräber von Sipán« (2000/2001), die kunsthistorische Ausstellung »Venezia – Kunst aus Venezianischen Palästen Sammlungsgeschichte Venedigs« (2002/2003) sowie die Ausstellung »Schönheit und Seele Japans« (2003) evaluiert. Im RLM wurden Evaluationen für die Ausstellung »Roots – Wurzeln der Menschheit« (2005/2006), die Ausstellung »Krieg und Frieden – Kelten, Römer, Germanen« (2007) und die Ausstellung »Alter in der Antike – Die

¹ Dieser Artikel ist Dr. Hanns-Ulrich Mette († 2004), Leiter der pädagogischen Abteilung der KAH, gewidmet, mit dem das ZEM in enger und freundschaftlicher Kooperation die erwähnten KAH-Projekte durchführen durfte.

Blüte des Alters aber ist die Weisheit?»(2008/2009) durchgeführt. Das RLM wurde außerdem nach einem umfassenden Umbau und der Neueröffnung im Jahr 2004 als kulturelle Einrichtung insgesamt evaluiert.

Der überwiegende Teil dieser Evaluationen war dabei als ex-post-Evaluationen konzipiert, d.h. die jeweilige Ausstellung wurde in abschließender Weise im Hinblick auf ihren Erfolg bewertet. Im Unterschied zur traditionellen Definition von Erfolg, die sich vorwiegend an den erreichten Besucherzahlen orientiert, hat das ZEM dabei den Blickwinkel in Richtung inhaltlicher Vermittlungsziele und einer Bewertung der Ausstellungsarchitektur erweitert. Auf diese Weise wird es auch kleineren Häusern mit eher geringeren Besucherzahlen ermöglicht, den Erfolg einer Ausstellung differenziert zu bewerten und zu dokumentieren.

Ein weiterer Untersuchungsansatz bezieht sich auf das Konzept der ex-ante-Evaluation – u.a. verwendet bei der »Venezia«- und der »Roots«- Ausstellung –, bei dem eine in der Konzeptions- und Planungsphase befindliche Ausstellung vor ihrer Realisierung »systematisch auf die Besuchergerechtigkeit hin durchgetestet wird« (AfeB e.V., 2003). Konkret umgesetzt kann das, wie z.B. bei der »Venezia« Ausstellung, bedeuten, dass vorab ermittelte Besucherinteressen hinsichtlich spezifischer Fragestellungen, z.B. Tourismus oder Hochwasser in Venedig, in die Konzeption der Ausstellung mit einfließen. Der Erfolg derartiger Maßnahmen des Qualitätsmanagements wurde dann – wie im Fall von »Venezia« – mit Hilfe einer gesonderten ex-post-Evaluation überprüft.

Ziel der Evaluationsprojekte

Alle Evaluationsprojekte wurden durch angehende Diplom-Psychologinnen und Diplom-Psychologen des Instituts für Psychologie der Universität Bonn unentgeltlich in Gruppenarbeit und unter Anleitung des ZEM durchgeführt. Die Studierenden erhielten durch diese Projekte – als Ergänzung ihrer theoretischen Ausbildung an der Hochschule – Einblicke in die praktische Durchführung einer wissenschaftlichen Evaluation. Wissenschaftliche Evaluation wurde dabei, entsprechend der Definition von Wottawa und Thierau (1998), verstanden als »Planungs- und Entscheidungshilfe«, »ziel- und zweckorientiert« sowie »am aktuellen Stand wissenschaftlicher Techniken und Forschungsmethoden angepasst«.

Die Zielsetzung dieser Evaluationsprojekte war dabei nicht ausschließlich, den Erfolg oder Misserfolg der jeweiligen Ausstellung festzustellen oder eine anschließende Rechtfertigung über den Einsatz finanzieller und personeller Ressourcen zu erhalten. Die Evaluierung der Ausstellungen wurde vielmehr mit einem ganzheitlichen, auf die Durchführungsqualität abzielenden Ansatz vorgenommen. Das Interesse eines Museums als Auftraggeber an einer Evaluation lässt sich am treffendsten mit einer Äußerung des Leiters des

Fachgebiets Pädagogik der KAH, Dr. Hanns-Ulrich Mette, zusammenfassen: Externe Evaluation wird demnach vor allem deswegen als fruchtbar beurteilt, weil in einem »[...]Ausstellungshaus mit einer hohen Frequenz von ständig wechselnden und sich überlappenden Ausstellungen ein professionelles Defizit besteht, dass sehr häufig keinerlei Zeit für eine kritische (Selbst-)Reflexion der eigenen Arbeit sowohl in ihren problematischen als auch erfolgreichen Aspekten bleibt« (Rietz, Mette & Kleinemas, 2003). Entlang dieses Grundgedankens wird wissenschaftliche Evaluation als ein wertvolles Instrument betrachtet, das es erlaubt, die eigene Arbeit objektiv und konstruktiv zu bewerten und Optimierungspotenziale zu identifizieren. Die Fokussierung liegt daher nicht auf der Beschreibung von Defiziten, sondern vielmehr im Aufzeigen von Entwicklungsmöglichkeiten.

Im Folgenden soll der Ablauf einer wissenschaftlichen Evaluation in einem Museum exemplarisch an der 2004 durchgeführten Interim-Evaluation des neu eröffneten RLM umrissen werden. Im Gegensatz zu den beiden oben beschriebenen Evaluationsformen (*ex-ante* und *ex-post*) wird die Interim-Evaluation *während* der laufenden Ausstellung durchgeführt und zielt nicht auf eine endgültige Bewertung ab, sondern ist eher im Sinne eines Monitoring zu verstehen, wodurch den Entscheidungsträgern eine Optimierung der Zielerreichung ermöglicht wird (Hong & Mark, 2003).

Ein exemplarisches Projekt

Nach längeren Umbauarbeiten öffnete das RLM im November 2003 erneut seine Türen für die Besucher. Einhergehend mit dem Einsatz moderner Architektur war auch eine inhaltlich museumspädagogische Neuorientierung angestrebt worden. Im Hinblick darauf war es der Museumsleitung wichtig festzustellen, wie das RLM nach dem Umbau von den Besuchern wahrgenommen und erlebt wird.

Im Rahmen einer Lehrveranstaltung wurden die Studierenden zunächst theoretisch auf die Durchführung des Evaluationsprojekts vorbereitet. In einem ersten Treffen machten die Kuratoren des RLM deutlich, wo bei dieser Evaluation ihre groben Interessenschwerpunkte lagen: Wie gefällt den Besuchern die neue Architektur? Wie finden sie die themenorientierte Struktur der Ausstellung? Wie gefallen den Besuchern die Angebote, die das RLM ihnen macht? Wird das Medienangebot genutzt? Wie werden die Werbemaßnahmen wahrgenommen und bewertet? Welche Möglichkeiten gibt es, mehr Besucher für das Museum zu gewinnen? Gibt es bezüglich dieser Fragestellungen Unterschiede bzw. Schnittmengen zwischen den Museumsbesuchern verschiedenen Alters? Was muss unternommen werden, um bestimmte Zielgruppen vermehrt für einen Besuch im RLM zu gewinnen?

In weiteren Treffen mit den Museumsvertretern wurden aus diesen globalen Zielsetzungen und Fragestellungen konkrete Evaluationsschwerpunkte

herausgearbeitet und formuliert. Es ergaben sich drei Themenbereiche – Rahmenbedingungen des RLM, Angebote des RLM und Image des RLM –, die mittels eines Fragebogens untersucht werden sollten. Unter diesen übergeordneten Themenbereichen konnten alle erarbeiteten Interessenschwerpunkte subsumiert werden. In enger Abstimmung und Zusammenarbeit mit den Kuratoren wurden daraus Fragenentwürfe formuliert und in weiteren Arbeitsschritten modifiziert und verdichtet. In einem Pretest im RLM wurde die Praxistauglichkeit des Befragungsinstruments sichergestellt.

Die Erhebung der Daten begann Anfang April 2005 und endete Anfang Juni 2005. Die Fragebögen wurden an Wochentagen vom Personal des RLM an die Besucher ausgegeben, am Wochenende waren zusätzlich Studierende der Universität Bonn vor Ort. Die Resonanz der befragten oder beobachteten Museumsbesucher auf die Evaluation war nahezu durchgängig positiv. Es war festzustellen, dass die Museumsbesucher sich als »Experten« ernst genommen fühlten und erfreut waren, dass ihrer Meinung Beachtung geschenkt wurde – dies – es sei noch einmal betont – ist bei allen Museumsevaluationen des ZEM eine durchgängige Erfahrung.

Wie die Ergebnisse der Interim-Evaluation im RLM zeigten, hat das RLM in den Augen der Besucher nach der Neueröffnung und durch den Umbau nicht an Attraktivität verloren: Durch die moderne Architektur und das themenorientierte Ausstellungskonzept wurden die älteren Menschen als Besucher gehalten und gleichzeitig das Potenzial geschaffen, auch junge Erwachsene, Kinder und Jugendliche mit einer gelungenen Mischung von Wissen und Unterhaltung zum Besuch des RLM zu animieren. Die Erwartungen der Mehrheit der Museumsbesucherinnen und -besucher wurden erfüllt, in vielen Fällen sogar übertroffen. Das neue Konzept der themenorientierten Ausstellungsstruktur wurde von der Mehrheit der Befragten wahrgenommen, dennoch besichtigten sie das Museum überwiegend systematisch und die unteren Ausstellungsräume vor den darüber liegenden. Das neue und erweiterte multimediale Angebot, z.B. Audioguides oder Informationsterminals, wurde sehr gut aufgenommen und stellte eine wichtige kostenlose Ergänzung und interaktive Bereicherung der Ausstellung dar. Besucher aller Altersgruppen hatten dabei kaum Nutzungsvorbehalte bezüglich möglicher Handhabungs- oder Funktionsschwierigkeiten; eine vermehrte Nutzung durch ältere Besucher mittels einer deutlicheren Herausstellung einer generationsübergreifenden Nutzungsintention wäre dabei anzustreben. Wie bei der neuen Konzeption des Museums intendiert, waren Unterschiede zwischen den Altersgruppen in Hinblick auf die Wahrnehmung und Beurteilung des RLM nur sehr vereinzelt festzustellen.

Ähnlich wie in vielen anderen Museen ließ sich feststellen, dass das RLM geringere Besucherzahlen hatte als die positiven Beurteilungen vermuten ließen. Nur eine Minderheit der Besucher wurde durch die durchgeführten Werbe- und Marketingmaßnahmen, z.B. Plakate an U-Bahn Haltestellen, auf

das RLM aufmerksam. Stattdessen wirkten vor allem Medienberichterstattungen und Mundpropaganda. Entsprechendes findet sich auch in einer Untersuchung von Schäfer (1996), hier nennen 30% der Befragten die Mund-zu-Mund-Propaganda als maßgeblich für den Besuch im Haus der Geschichte in Bonn (vgl. auch Adams, 1996). Ältere Besucher fanden den Weg ins RLM meist ohne gezielte Werbemaßnahmen, sie informierten sich überwiegend eigeninitiativ. Um in gleichem Maß Besucher jüngeren Alters für das RLM zu gewinnen, wurde empfohlen, die Werbemaßnahmen speziell auf diese Altersgruppe auszurichten, beispielsweise durch die Schaltung von Anzeigen im Internet (vgl. Schäfer, 2004).

Die Evaluation zeigte, dass dem RLM von den Befragten fast durchgängig ein sehr gutes Image bescheinigt wurde und dem Museum ein hoher Stellenwert in der Bonner Museumslandschaft zugeschrieben wurde. Es war daher zu resümieren, dass eine Ausweitung der Werbemaßnahmen zu einer Erhöhung der Besucherzahlen führen sollte.

Die Präsentation der Evaluationsergebnisse für das weite Spektrum der hausinternen und externen Interessengruppen erfolgte in einer öffentlichen Veranstaltung im RLM, wobei die Ergebnisse des umfangreichen schriftlichen Evaluationsberichts in zusammengefasster Form als Management-Report vorgestellt wurden. Analog geschah dies auch immer bei den Evaluationen in der KAH, nicht zuletzt auf das Betreiben und im Interesse von Dr. Hanns-Ulrich Mette im Auditorium der KAH – für die Studierenden ein Musterbeispiel von integrativer Vermittlung von *soft skills* oder *business skills*, für die »Auftraggeber« der Beginn der nächsten Iteration im Evaluationszyklus.

Fazit

Die kontinuierliche Zusammenarbeit des ZEM mit Bonner Museen, prototypisch für die Kooperation im Bereich der Evaluation zwischen theoretischer Wissenschaft und angewandter Praxis, kann für alle Beteiligten als gewinnbringend betrachtet werden: Das ZEM kann seine Kompetenzen im Bereich der Evaluation kultureller Einrichtungen festigen und ausbauen, die teilnehmenden Studierenden erhalten praktische Einblicke in das Berufsfeld des Evaluators, und die Museen erhalten kostenneutral qualitativ hochwertige Evaluationsergebnisse zur weiteren Optimierung ihrer Arbeit, was nicht zuletzt auch ihren Museumsbesucherinnen und -besuchern zugute kommt. Dies setzt allerdings auf Seiten des Auftraggebers Aufgeschlossenheit für Möglichkeiten und Grenzen der wissenschaftlichen Evaluation voraus sowie die grundsätzliche Bereitschaft, aus den Ergebnissen abgeleitete Handlungsempfehlungen auch tatsächlich umzusetzen. Es kann als besonderer Erfolg der bisherigen Kooperationen zwischen ZEM und Museen betrachtet werden, dass auf der Basis zielgerichteter Austausch- und Kommunikationsprozesse

zwischen den Partnern in fast allen Fällen Umsetzungsmaßnahmen auf die Evaluation folgten, die zu einer messbaren Qualitätssteigerung der Ausstellungsangebote führten.

Literatur

- ADAMS, G. D. (1996). Museumsmarketing und Besucherentwicklung. In Haus der Geschichte der Bundesrepublik Deutschland (Hrsg.), *Museen und ihre Besucher* (S. 165-175). Berlin: Argon.
- ARBEITSGRUPPE für empirische Bildungsforschung (AfeB) e.V. (Hrsg.). (2003). *Einführung in die Evaluation von Museen und Ausstellungen*. (Teil des Lehrbriefs. Evaluation und Besucheranalysen. Empirische Informationsgewinnung in Museen und Ausstellungseinrichtungen des Instituts für Kulturmanagement an der Fernuniversität Hagen).
- FRIEDMAN, A. J. (1996). Why museums don't evaluate. *Visitor Behavior*, 11 (2), 6-8.
- HONG, H. D. & Boden. M. (2003). *R&D Programme Evaluation: Theory and Practice*. Ashgate Publishing Ltd.
- RIETZ, C., Mette, H.-U. & Kleinemas, U. (2003). Evaluation von Ausstellungen in der Kunst- und Ausstellungshalle der Bundesrepublik Deutschland in Bonn: Alle gewinnen! *Museen im Rheinland*, 2/03, 12-14.
- SCHÄFER, H. (2004) Besucherforschung als Basis für neue Wege der Besucherforschung, In B. Commandeur. & D. Dennert (Hrsg.). *Event zieht – Inhalt bindet. Besucherorientierung von Museen auf neuen Wegen*, Bielefeld 2004, S. 103-119.
- SCHÄFER, H. (1996). Besucherforschung im Haus der Geschichte. In Haus der Geschichte der Bundesrepublik Deutschland (Hrsg.), *Museen und ihre Besucher* (S. 143-155). Berlin: Argon.
- WOTTAWA, H. & Thierau, H. (1998). *Lehrbuch Evaluation* (2., vollst. überarb. Aufl.) Bern: Hans Huber.

Georg Rudinger, Kieran Donaghy & Stefan
Poppelreuter

Societal trends, mobility behaviour and sustainable transport in Europe and North America¹

Abstract

This contribution describes the work of Focus Group three of the European Union network Sustainable Transport in Europe and Links and Liaisons to America (STELLA). It examines especially social and behavioural aspects of sustainable transport from a transatlantic perspective. Significant societal trends (e.g. the ageing of societies) are surveyed and their implications for mobility behaviour are drawn. The sustainability of this behaviour is considered along with constraints and drivers of this behaviour in Europe and North America. The contribution takes up relevant policy issues and concludes with a discussion of a transatlantic research agenda on social and behavioural aspects of sustainable transport.

1. Introduction: The STELLA network and its Focus Groups

The sustainability of existing and emerging patterns of mobility behaviour in Europe and North America is questionable and is a subject of grave concern for transportation researchers, practitioners, and public authorities on both sides of the Atlantic. Many of these concerned individuals have been participating in the three year (2002-2005) STELLA Thematic Network project. STELLA is an acronym for Sustainable Transport in Europe and Links and Liaisons with America. The STELLA project is part of the European Commission's 5th Framework Programme for Research and Development and is centred on common issues in Transatlantic transport research. STELLA aims to generate value added from knowledge exchange and to support a common research approach from both sides of the Atlantic, which will not only be to the benefit of the research community, but also of interest to policymaking bodies and industrial organisations. North Americans participated in the project through the STAR network (Sustainable Transportation Analysis and Research), which is supported by the US National Science Foundation, and Transport Canada.

1 Presentation of FG 3 »Society, Behaviour and Public/Private Transport« NECTAR conference – STELLA Meeting, Las Palmas, June 2nd – June 4th 2005

The STELLA network has been addressing five major focus areas which have been identified as critical fields of interest for a transatlantic thematic network in the transportation field through dedicated focus groups. The STELLA thematic network has prompted an intense debate in both Europe and North America on the future perspectives of transport research and the foundations of policy analysis in this field (including the question on the formulation of policy lessons). Focus group members hail from the academy, industry, and government. Focus Groups 1 and 2 address the themes of Globalization, E-Economy, and Trade; and ICT, Innovation, and the Transport System. Focus Groups 4 and 5 address the themes of Environment, Safety, Health, Land Use, and Congestion; and Institutions, Regulations, and Markets in Transportation.

Focus Group 3 has been examining the dynamics between Society, Behaviour, and Private/Public Transport. Focus Group 3 has met three times, in Bonn in April 2002, in Arlington, Virginia in January of 2003, and in Lisbon in May, 2004. The principal objectives of these meetings have been to explore the interrelationship between social change and transport in the different spatial organisations and cultural contexts of Europe and North America, to identify commonalities and contrasts in forces determining spatial mobility in Transatlantic societies, to mobilise high quality research, and prepare central findings on the relationships between society, behaviour, and private/public transport for dissemination to a wider international audience.

2. Societal Trends and implications for mobility behaviour

Social change affects transport and is affected by transport. In close interaction with social values and individual attitudes, mobility and transport patterns are linked in complex ways with significant social trends, such as the adoption of suburban life-styles, the ageing of populations, the decline of the nuclear family, etc. It is increasingly evident that the corresponding (modern) lifestyles in affluent societies, and the mobility behaviours associated with such life styles, are not consistent with protection of environmental quality, efficient use of human, natural, and financial resources, and promotion of social cohesion and just distributions of opportunities and costs of using transport systems. Arguably, whether or not specific societal trends or combinations of trends are sustainable may depend in large part upon whether or not changes in individual behaviours can be brought about through policies that affect drivers and constraints upon relevant types of action.

This paper examines social and behavioural aspects of sustainable transport from a transatlantic perspective. In so doing, it draws heavily on papers that have been contributed to the above mentioned meetings of STELLA Focus Group 3 (Society, Behaviour, and Public/Private Transport) published in Donaghy, Poppelreuter & Rudinger (2005).

In the 20th and 21st centuries, both Europe and North America have witnessed urban settlement patterns of outward growth and lower densities and increased demands for personal transportation. As locations of residential, work, and leisure activities have become increasingly separated spatially, mobility has increased and come to be valued more. Social networks have become more far flung and there are perhaps fewer overlapping uses in the same place. Industrial production and distribution has also become more transit intensive as the relative cost of transport in production has steadily decreased. We consider first the broad sweep of these changes and then specific details of the specific cases of the United States, countries of the European Union, and Central and Eastern Europe countries.

As Axhausen (2005) has observed, the key societal development of the last 100 years has been the translation of accumulated differences between productivity and population growth into mass consumption. The key sectors in this process have been housing and transportation. In the period of reference, most populations of the OECD countries have achieved almost complete motorization. Increased access to transport has been translated into use and associated speed gains translated into longer distances travelled. The literature suggests consumers have a constant »travel budget«, or tend to spend the same amount of time travelling regardless of speed of travel, and the pattern is discernible in all OECD countries (Axhausen, 2005).

Increases in housing consumption, distances travelled, and telephone contact have led to a rearrangement of expenditures from increasing household budgets. While expenditure shares of food have dropped considerably, expenditure shares of leisure, education, travel, and communication have all nearly doubled since the early 1900s. While the larger shares of expenditures on leisure and education reflect longer life expectancy, more years spent in education, shorter work weeks, and longer retirements, the adoption of the automobile is the primary reason for the increase in the expenditure shares of travel and communication.

The motorized, connected, and spatially dispersed populations of today have different patterns of work and socialization than the populations of 50 years ago. Daily life revolves around friends, family, work, school, and shopping. The distribution over space of these commitments and opportunities shapes the activity space of a person in his or her everyday environment. The size of this activity space determines a person's consumption of transport services. Increasing distances between the poles of a person's activity space have far reaching and permanent effects, which cannot be reversed quickly, given long-term investments made in work, family, and friendship. Society has increasingly become based on spatially dispersed networks, which are weakly overlapping. The breakdown of walking environments and the weakening of closed social milieus have been contributing factors. In view of these trends, one may assume that not only will the members of social net-

works be spatially dispersed, but social networks will be less coherent – i.e., fewer people who are spatially proximate will share multiple affiliations.

Within the last 25 years personal miles travelled in the United States (U.S.) has increased by 80%, a rate which is almost 3 times that of population growth over the same period. While there has been increased demand for automobile use over the last few decades, the same cannot be said for bus and transit use. The downward trend in transit ridership, which began in the 1950s, can be explained in part by the patterns of metropolitan form that have evolved (Schintler, 2005). In many metropolitan areas, the suburbanization process has given rise to polycentric patterns of development and the emergence of edge cities, urban sprawl, and areas of low-density development. These patterns are not conducive to transit use, neither bus nor rail, as they lack the density required to support efficient operation. The development of information and communications technologies and their use in control of production and supply-chain logistics have contributed to a functional division of labour in space or vertical and horizontal fragmentation of industrial production, the development of just-in-time and other approaches to production that are generally more transport-intensive. In the United States the logistics revolution has contributed significantly to the volume of freight shipments. Commodity-flow data for 1993 and 1997 indicate a 3% increase in the value of shipments but a 27% increase in tonnage and a 12% increase in ton-miles (Munroe & Hewings, 2007). Most trade in both North America and Europe is now *intra*industrial in nature and interregional trade is increasing because of increased specialization of establishments and increased complexity of production processes. Economies of scope, as well as scale, are being exploited at the multi-plant level (Jones & Kierzkowski, 2001).

In the countries of the European Union (prior to May 2004) there has also been a spectacular increase in automobile use. Presently 80% of all passenger journeys are made by car, 8% by bus, 6% by rail, and 5% by air (European Commission, 2003). Air transport has grown rapidly, with passenger traffic rising at an average annual rate of more than 7% over the last 20 years. Rail transport, however, has been in decline as market share for passenger traffic has declined from 10% to 6%, and from 21% to 8% for freight. This decline reflects some problems of rail: railways are slower and less reliable than road transport, national railways operate according to different standards and have not done enough to integrate with each other. While water transport is viewed as the poor relation among different forms of transport, 41% of goods transported within the EU travel by ship over short-sea routes.

Suchorzewski (2005) reports that, in the Central and Eastern European Countries (CEEC), since 1990 passenger trips made by car have increased by 5 to 7% per year. The number of passengers served by railways has dropped by about 5.7% per year, as has use of intercity bus transport. In some CEEC the state has completely withdrawn from its role as owner and financier of

urban transport, while still preserving part of control over urban finances. In some countries, such as Poland, emphasis has been placed on mobility management in an effort to decouple economic growth from transport intensity.

Although walking and cycling are the most economically, socially, and environmentally sustainable forms of human mobility, they are not adequately catered for in policies or practice in either North America or Europe and are in decline (Gaffron, 2005). We note that in the European Union there is a significant unfulfilled potential for them since 50% of all trips are shorter than 3 kilometers, a distance which can be walked in 25 minutes or cycled in 10. Yet statistics show that the modal share of walking and cycling in Europe is declining (DGTREN, 2002). Spatial spread in social networks, motorization of private transport, and policies encouraging more of the same have together contributed further to a decline in walking and cycling.

2.1 Mobility needs of the socially excluded, elderly, and female transport system users

The nature and pattern of social exclusion is quite different between the U.S. and European countries, as well as among European countries. First, redistribution policies are more limited in the U.S. Second, social exclusion in the U.S. is heavily concentrated in the inner city and rural areas, mainly due to the intersection of poverty and exclusion with race and ethnicity and related discriminatory practices in the U.S. housing market (Kaplan & Halloway, 1998). Third, the dominance of the private vehicle in U.S. urban transport makes automobile mobility essential everywhere, with the possible exception of the centres of the largest cities. Low-income households are much less likely to own an automobile, and, when they do, it tends to be much older. Low-income individuals tend to walk and ride the bus more than higher-income individuals.

Certain groups are disproportionately at risk of social exclusion and people are most vulnerable at particular stages of the life cycle. The primary dimension of social exclusion is poverty. Of those suffering relative poverty, six groups are most at risk: children, young people, older women, the unemployed, single parents and large families.

Because the U.S. poverty definition is based on fundamental needs (e.g. food consumption requirements), adjusted for household size and composition, the poverty rate defined is much lower than in the EU. The European Council's joint report on social exclusion, published in December 2001, recognizes the need for a broader approach assessing exclusion, which not only emphasizes low-income but also access to equally relevant aspects such as jobs, education, housing, healthcare, the degree of satisfaction of basic

needs and the ability to participate fully in society (Council of the European Union, 2001).

Pickup & Giuliano (2005) have identified transport related factors that compound the primary risks of social exclusion in both Europe and the U.S. Several examples are as follows. Travel patterns become tied to the monthly cycle of social benefit payments, contributing to access and mobility patterns of the unemployed. Difficulties accessing facilities limit the ability to maintain social contact of low-income elderly people who are reliant on public transport. The high cost of transport for low-wage workers restricts their social activity space, reinforces a local lifestyle, and increases the likelihood of further social exclusion. Children and teenagers from transport disadvantaged families suffer from their limited ability to travel beyond their »disadvantaged« area.

The populations of both continents have been ageing as the proportion of the population at 65 years of age or older is at its historical high. People are living longer and having fewer children, although the latter phenomenon is more pronounced in Europe (Economist, 2002.) Of those living longer, more are women. With the aging and feminization of the population has also come a shrinking of the middle class and an increased spatial concentration of low-income members of society (see, e.g. Krugman, 2002).

In both Europe and North America women's travel patterns are converging with those of men at the aggregate level, in the sense of increased automobile use and less use of transit, walking and cycling. Where women's travel patterns differ from those of men is at the micro level. Whereas trips of men tend to be more single-purpose, women's trips chain more (Schintler, 2005). Kwan's (2003) research concerning personal spatial-time prisms also reveals that women tend to be more constrained in their mobility by their obligations as primary care providers. Such obligations also contribute to women having more complicated and difficult commutes than men. As more women drive, their accident rates are expected to increase proportionately, but injuries are expected to increase more than proportionately. This is so not because women are less safe drivers than men, but because they are more vulnerable physically (Rosenbloom & Hakamin-Blomqvist, 2004). Even as these generalizations can be made, it is important to note that women's travel needs and problems are not homogeneous. Elderly, female, and socially excluded (principally, low-income) users of public and private transport have different mobility needs which are met more or less well by existing systems.

2.2 Ageing and its implications for mobility behaviour

Especially aspects concerning the ageing of societies and elderly persons have been taken into account in Focus Group 3 (See also European Conference of Ministers of Transport, 2000). One of the most significant societal

trends is the ageing of societies. An increasingly larger proportion of the population is reaching old age. In nearly all countries of the western world the elderly already make up over 12% of the total population, and this percentage is projected to increase substantially year by year. An increasing number of these individuals are licensed to drive, and they drive more than their age cohorts a decade ago. In view of the constantly increasing number of old and very old persons in our societies it is no surprise that their mobility behaviour has become an issue of public and scientific interest (Schaie & Mollenkopf, 2005). On the one hand, mobility (the ability to move about) and traffic (the transportation of people, goods, and news) have become an even more important precondition of ensuring the ability to lead one's everyday life, keep up social relations, take part in every kind of activity outside one's own four walls, and seek out places subjectively significant or objectively central to provide for daily material needs and health care. On the other, mobility is increasingly jeopardised as a person ages (Waldorf & Pitfield, 2005). The profound changes in the demographic structure due to the growth of ageing populations in developed countries, especially in the European Union countries and the United States, will clearly require a profound reorganisation of society, both in public services and in private supply. The elderly users will be qualitatively different, as they will work longer, be more alone, and have a higher level of education. The majority of the elderly would probably require greater mobility than at present, regarding the need to work, travel and have interpersonal contacts (Schaie & Mollenkopf, 2005). However, it can also be assumed that the current ›ageing of ageing‹ will increase the number of persons with reduced mobility performance or with cognitive and sensory disorders (e.g. sight, hearing), who today are often only potential users or excluded ones. Therefore emphasis should be placed on the expected increase in the number of elderly persons with mobility impairment and the transportation challenges associated with this change. In order to keep older persons mobile and allow them to be independent it is important to know their mobility patterns, attitudes to transport and different needs, with special attention given to elderly persons with mobility impairment.

For the elderly, mobility is a quality of life issue. It is essential for maintaining attributes of personal choice and independence, familiar habits, and lifestyle. Mobile elderly are able to reduce personal isolation, participate in recreational activities, obtain goods and services, maintain access to financial and personal consultants, remain active in a religious or spiritual community, and maintain intellectual stimulation. Although ageing is accompanied by waning strength, the needs of the elderly and disabled are different.

While the populations of all European and North American countries are ageing, the ageing process differs across countries. Simultaneously, there is a feminization process (or increase in percentage of women in the population)

that varies across countries. As women age they tend to drive less or cease driving altogether more so than men. The growing difference between the numbers of elderly who survive and cease to drive (or do not drive) and those who survive and continue to drive indicates that there will be a growing demand for trips that are not auto-dependent (Tacken, 2004). While the difference is greater for women than for men, the number of elderly men and women needing public transportation services will be increasing. Based on multiple-risk survival models Waldorf & Pitfield (2005) predict that this difference will become particularly critical in the U.S. around the year 2018. This potential development has a number of implications.

So, different perspectives can be chosen to analyse the relationships between influencing factors and elderly people's outdoor mobility and traffic behaviour in general. One focuses on ›causal‹ factors. For example, what socio-demographic and psychological characteristics are associated with the driving behaviour of the elderly? This perspective is chosen frequently in traffic research studies. Another perspective might be called a segment-specific approach: Can the elderly as a group be subdivided into homogeneous subgroups? For example, are there variables that characterise those elderly persons who have a positive attitude toward public transport?

It is important to bear this second perspective in mind, since ›the elderly‹ – and particularly the ›very old‹ – are a heterogeneous group (Schaie & Mollenkopf, 2005). This helps us to identify and understand their different mobility needs, their diverse driving behaviour, and the variety of their compensatory strategies (Rudinger & Jansen, 2003). This perspective allows to determine the size of the different homogeneous subgroups of the elderly and to identify the characteristics of these subgroups as a basis for improving traffic safety and options for outdoor mobility. Based on life-style differences (e.g. stimulation-seeking, intellectually curious, indifferent, passive, and negativistic) it is possible to isolate different types of mobility patterns of the elderly (Rudinger, 2002). Typological approaches confirm the existence of groups of elderly deserving special attention with regard to intervention, rehabilitation and prevention (Willis, 2003).

In addition to a need for greater provision of public transportation, especially if older persons should be convinced to quit driving, traffic management will need to be adapted to the changing driving population (Hakamies-Blomqvist, 1993; Rosenbloom & Stahl, 2002). Road markings and street signs will need to be made more visible, intersections simplified, HOV lanes increased and infrastructure and vehicles redesigned. The relocation of residences and facilities for retirement communities will also need to be better planned. Public transportation will need to adapt to the needs of elderly users and information technology used more widely. And if ›ageing in place‹ is to become a real option for more elderly men and women, more demand-responsive transit will need to be provided. Nevertheless, the car will be

widely used by the older population, and car industries will therefore have to design cars that meet the real needs of the elderly and their gradually declining functions. In this direction the new technologies (ICT, telematics) might give considerable support.

A comprehensive notion of mobility should take into account the degree of mobility realised, the range of activities engaged in, the diversity of transportation modes available to users, and satisfaction with mobility.

2.3 Society, behaviour and sustainable transport

Against this background of information on trends in settlement patterns, mobility behaviour, and the habits and needs of specific user groups (e.g. the elderly), we consider meanings of sustainable transport. Virtually all notions of sustainability have to do not only with the persistence of systems but also the conveying of a healthy state of a system from one generation to the next. Since the Brundtland Report (World Commission on Environment and Development, 1987) it has been recognized that sustainability must be viewed in a number of dimensions. Minimally these would include environmental, economic, and social (Heal, 1993). But there is also an argument for the inclusion of the cultural dimension. The question of sustainability *vis-à-vis* transport systems concerns not only what level of personal mobility and flexibility is environmentally and economically sustainable (for living and future generations) but also what differences in transport service availability and effective access can be consistent with social cohesion. Lakshmanan (2001) has identified five properties of a sustainable transportation system: i) urban forms must be planning driven and transit-friendly; ii) transit mode choices should be market driven; iii) infrastructural investments should be efficiency driven; iv) financing must be innovative; and iv) the system must be flexible enough so that it can be reinvented. With this dimension in mind we might also ask what level of environmental quality is *socially* sustainable.

If our aim is to arrive at settlement patterns, transport systems and patterns of usage that are sustainable, we need to understand better the constraints upon and drivers of mobility behaviour. In a recent survey of transportation and land-use linkages, Medda & Boarnet (2003) have identified two types of linkage. In the first, transportation infrastructure influences land use, and in the second, urban form affects travel behaviour. Empirical studies reveal that in the former, the speed of adjustment of land use to transportation infrastructure may be measured in decades, whereas in the latter, the speed of adjustment of behaviour to urban form may be instantaneous. Among the crucial findings made are that mixed land uses are associated with a reduction in trips, the growth of road networks coincides with population growth, and the influence of transportation infrastructure on development is large. In an elaboration of analytical models (based on fairly strong assumptions) Medda

& Boarnet (2003) demonstrate how transport costs can influence the length and width of a central business district (CBD), how the area allocated to transport can influence the shape of CBD, and how urban form is affected by traffic flows of different transportation modes. The models they consider may be used to analyze how the amount of area dedicated to different transportation modes will affect urban form.

In a recent survey of literature on behavioural aspects of road users Stern & Richardson (2005) have examined formal and empirical characterizations of travel behaviour. They trace a progression of behavioural modelling approaches from random-utility to rule-based paradigms and activity-based and process-oriented frameworks. In the latter type of model, motives and needs drive decisions made under constraints, which lead to response behaviour (in the form of driving behaviour and travel demand) over space and time. In comparative studies of travel patterns in Europe and North America, differences in travel and mobility can be traced to constraints that may take the form of income, network densities, transport technologies, and social trends. Since models are conceptual in nature, determining their usefulness in understanding behavioural aspects of transport behaviour requires empirical testing. Appropriate data, however, are often difficult to obtain and hence much about travel behaviour remains unaccounted for empirically. Stern & Richardson (2005) propose the use of decision-field theory, in which the process of deliberation is explicitly considered.

In considering questions that empirical studies have framed, Stern & Richardson (2005) observe that there are issues concerning long-term versus short-term decision making, socioeconomic determinants, and cultural differences. The residential decisions of multiple-occupation households would appear to be particularly important to investigate. Stern & Richardson (2005) observe that considerable research has been conducted on factors affecting the time of day of travel, route choice, optimal commuting patterns, modal choice, vehicle choice, and urban form. It may be the case, however, that multiple decision processes are involved in the determination of travel behaviour and that we are unlikely to get very far in policy making without better theories of what drives travel behaviour in the context of multiple decisions. From the perspective of empirical research, it is easier to test specific hypotheses about travel behaviour than to operationalize aggregate models in which compound hypotheses are empirically confronted. Stern & Richardson (2005) also note that even if improvements in theoretical explanations of travel behaviour can be achieved, it remains to translate them into effective policy prescriptions. And it is questionable whether a more universal model of transportation behaviour would explain observed differences when people in different situations often use different decision-making heuristics.

Current transport system configurations and their use have come about through a combination of technological innovations, private initiatives, and

public policies. There is a certain social inertia built into transport systems that is difficult to change. And even as we may come to understand better the mobility needs and behaviours of particular user groups or individuals, understanding individual behaviour may not help us to understand system effects that can be counter-intuitive (Vickerman, 2001.).

3. Policy considerations

Policies are commonly viewed as rules for responding to recurring conditions or situations. As such they can be both explicit and implicit about what they permit or forbid. Policies are often shaped in response to unmet needs. And policies may be used to develop information from which to learn about system responses (Intergovernmental Panel on Climate Change, 2001). Much we have discussed above has to do with the special needs of particular transport system users that are now unmet. Those needs are themselves conditioned by infrastructure provisions and long-term commitments.

A common distinction made between policies whose intent is to influence or modify behaviour is between those that are market-incentive based and those that are command-and-control in nature. In the area of transport policy, Europeans are more comfortable with both command-and-control and price measures (road pricing, fuel taxes) to recover costs of providing public transit and achieve reductions in congestion and emissions than are North Americans.

Policies are generally evaluated in terms of efficiency and equity. There are external costs to transportation behaviour not paid for by transport system users. These include economic, regulatory, and infrastructural costs. Any proper assessment of the equity of impacts must take into account how these costs (and the benefits derived from transportation systems) are distributed both intra- and inter-generationally and how transport facilities and services are distributed spatially. Economic theory suggests how optimal tax and investment rules should be constructed: The efficiency and equity of any policy reforms should be assessed in terms of whether or not the reforms are revenue neutral. To design better schemes for redistributing costs and benefits of transport systems, instruments other than taxes should be considered, locational analyses of the effects of transport policies should be made, and more detailed analyses of impacts on difference income groups should be conducted. Bae & Mayeres (2003) observe that equity tends to be evaluated mostly in terms of income. Typically, the lower a transport system user's income is the greater is the percentage of his or her income expended on transportation. Since adopting strict efficiency rules for pricing transportation will likely impact the poor adversely, a pro-equity policy may involve a second- or third-best approach in which there is a redistribution of revenues collected from externality-related cost fees. A pro-equity transportation pol-

icy must also focus on demand-side assistance. In spite of policies intended to increase transit use, metropolitan areas in the U.S. have experienced an increase in vehicle miles travelled without an increase in transit's share. While technological developments in the automobile industry hold promise for improvements in energy efficiency and air quality, their costs and diffusion time render short-term equity impacts negligible. If pro-equity policies are desired by society, both command-and-control and market incentive measures might be used to good effect.

In designing policies that meet the needs of different user groups and are efficient and fair in their delivery, transportation policy researchers will have to learn to package policy recommendations with strategies of adaptation and work more closely with targeted stakeholder groups. In general policy must take into account that sustainable transport and mobility systems require the co-existence of several different mobility modes. Moreover concepts of mobility management must be fostered and tax policies should not unsuitably favour a singular mobility mode.

We need a policy that avoids discrimination and social exclusion (mobility as a ›basic right‹, e.g. access to mobility for everyone, reasonable tax policies), especially concerning the elderly. Social, urban and traffic planning should be coordinated on micro-, meso- and macro-levels (e.g. implementation of advisory boards of elderly in local/city councils). The accessibility, safety and security of different mobility systems (from an objective and a subjective attitudinal perspective) should be improved. On a transport dimension a balance between modes and means of transport is necessary, so that those who give up the use of a private car have available alternatives of good quality, without any discrimination. On a social dimension it must be assured that citizens are provided with a mobility system adequate for their needs and that no exclusion through price, or any other criteria, will be imposed on the base of economic or financial grounds.

4. A transatlantic research agenda

The present political environments in Europe and North America are very different, but the aggregate pictures of transport system use by citizens are similar and are converging (Schintler, 2005; Suchorzewski, 2005). Whereas the European Union seeks to implement a pan-European policy whose particulars are being worked out at the national level, there is less of a unified thrust in North American countries. And whereas European cities are more ready to implement aggressive command and control and market oriented policy measures, North American cities are more passive.

Bringing about planned behavioural changes that promote sustainable transport systems will require (at a minimum) deeper understanding of complex social phenomena, hence it will require research on the behavioural

foundations of such phenomena. Conditions for applied comparative research in both Europe and North America regarding behavioural motives, innovative strategies and policy assessment in the transportation sector with a view to sustainable transport should be fostered and created. We identified a set of research questions that might be grouped loosely according to whether they are of a more fundamental or more applied (policy-oriented) nature and conclude with reflections on what is needed to encourage fruitful international comparisons and explore transition paths to sustainable transport.

Research infrastructure and research policy issues are necessary to encourage international comparative studies and to facilitate international cooperation in the transatlantic study of social aspects of sustainable transport. Common surveys should be carried out in the United States and Europe and will provide useful comparative findings. For example, as the populations of Europe and the United States continue to age, transnational collaboration is essential to create environments that meet the needs of the elderly so that they can live their later years as active and valuable citizens of our communities (Marcellini & Allen, 2004). Particularly, the study on the communalities/differentiations of these effects in different spatial-economic and social contexts has been identified as a necessary area of research. Cross-cultural research is then envisaged in the design of a comparative analysis between Europe and North American. There is also a need for benchmark assessments of how well common models capture stylised facts of travel behaviour in Europe and North America. Hence different spatial levels – from theoretical, methodological and empirical viewpoints – might be analysed, by investigating concepts, data, and particularly the different speeds of ICT and transport dynamics and behavioural responses.

In this framework, there is a need to deepen, in both the European and North American contexts: (a) theoretical perspectives (such as the possible change of paradigms or modelling, e.g. in the sustainability concept, decay functions, behavioural models, forecast methods, and in general, in transport theory), (b) methodological issues (such as the exploration of new research methods, e.g. structural equations, multilevel models, self-organised criticality issues), (c) empirical evidence on the ongoing (un)sustainable transport-economic and social scenarios and (d) a strategy for inferring meaningful policy lessons in the context of transport policy research. These ingredients appear to be a common focus, with a view to the related policy analysis and the future elaboration of a policy research agenda.

4.1 Fundamental research questions

A number of fundamental research questions are raised by the increasingly dispersed nature of social networks. Does the increased selectivity of social networks improve the quality of life? Will reduced local networking lead to

more reliance on commercial or governmental services, especially in times of crisis? What are the implications for short-term travel behaviour of longer-term commitments to residential ownership, employment, automobile or seasonal travel pass ownership, and social network membership?

Reflections on societal trends in North America and Europe give rise to questions about the nature of attitudes toward sustainability and the nature of the relationship between attitudes and actual decisions. To what extent can attitudes toward and the valuation of mobility within highly mobile cultures be changed and how much individual mobility can be permitted in a sustainable transportation system? Related questions concern how lifestyles can be influenced and what the effect of established land use patterns on lifestyles and attitudes might be. An open question of historical importance concerns the extent to which Central and Eastern European countries might be able to avoid the negative consequences of the growth of mobility experienced by Western European and North American countries. In view of the inertia of established patterns of travel behaviour, it is important to inquire how effective in facilitating progress toward sustainability road pricing, teleworking and telecommuting programs, intelligent transportation systems, alternative transport arrangements, or innovative management schemes might be.

Focusing on the needs of particular user groups, there is a clear need to establish a consistent definition of social exclusion (in its many forms) and its relationship with mobility needs as accommodated by transport systems. There is also a need to study how socially excluded people cope with restricted mobility and a need to compare U.S. and European cities vis-à-vis job access and employment. Regarding the latter, a relevant research question is whether the greater centrality and spatial concentration of European cities, together with their better public transport, lead to less disparity in job access and less »spatial mismatch« between locations of residence and employment than occurs in the U.S.? Or is there a European analogue to the »permanent underclass« of North American societies? What are the distributional implications of differences in the cost of owning and operating an automobile in Europe and in North America? Are the European elderly ageing in place or are they moving to less costly and more remote areas?

Thus improvements in the mobility situation of elderly must focus as much on transport policy and socio-political measures as on appropriate urban development planning. It is important to create flexible, user-centred options for mobility that offer a genuine alternative to both the private automobile and traditional local public transport services, and provide for neighbourhoods that also respond to the needs and wishes of an ageing population. For the elderly, whose life space contracts with advancing age because of their inability to overcome environmental obstacles, it is crucial that the areas near their homes have readily accessible stores, medical and care ser-

vices, appropriate public transport and other facilities that will allow them to continue leading independent lives and being full members of society.

4.2 Applied research questions

Among the questions that arise concerning the linkage of research to policy formulations and applications for achieving sustainable transport systems are the following. To what extent do different user groups require special policy interventions and how should we weight trade-offs between accommodating special needs and transportation system efficiency costs? How much do we need to know about behavioural drivers in order to design effective policies (to control, e.g. congestion and emissions)? What kinds of policy measures would best facilitate travel by both elderly and young? How can the findings of empirical research about transportation behaviour be made available in formats useful to policy makers? How should income redistribution schemes be designed to promote greater equity among different transportation system user groups?

4.3 Research infrastructure to encourage international comparison

To facilitate international cooperation in the transatlantic study of social aspects of sustainable transport there is a need for the development and sharing of data sets. There is also a need for benchmark assessments of how well common models capture stylized facts of travel behaviour in Europe and North America.

4.4 Exploring transition paths to sustainable transport

Five properties of a sustainable transport system were enumerated above. These properties do not indicate which particular system arrangements are in the feasible set of arrangements, but may help us to identify candidates for feasibility. Just how transportation researchers should go about exploring the feasible set in interactions with policy makers and stakeholders is itself a question to be researched. It seems reasonable, though, that researchers will need frameworks for conducting experiments about making transitions to alternative steady states, frameworks which will integrate models of transportation, land use, industrial production and distribution. Clearly such integrated models must have temporal, spatial, and sectoral scales and scopes that permit adequate characterization of the complex of problems. They must, moreover, characterize faithfully processes of disequilibrium adjustment to be able to reproduce the development of historical problems and convey a realistic sense of time frames of these problems' resolution — i.e., they must capture lumpy changes in durably built environments and long-term commitments of transport systems users that such environments influence. Such

modelling frameworks must enable researchers to distribute costs and benefits to different stakeholders, many of whom are not yet living, if they are to support political deliberations. Because our knowledge is limited, our foresight imperfect, and our habits persistent, such modelling frameworks need to incorporate explicit treatments of learning and behavioural change. And because system interventions in the form of policy decisions take time to implement and often induce unintended consequences, our models need to help us to think carefully about what and how certain types of decisions are conditioned by others.

Assuming significant headway can be made in research to support deliberations about sustainable transport policies, there remains other important work for researchers concerned about such policies. Vickerman (2001) urges transportation researchers to raise public awareness of systems relations by doing a better job of »connecting the dots« in public forums. He also urges researchers to work to persuade the public of the importance of and need for significant changes that many are reluctant to consider. Lakshmanan (2001) remarks upon the need to improve the quality and meaningfulness of interactions between researchers and stakeholders and the need to study how to package strategies for sustainable transport for implementation at the local and regional levels. The position of transport and communication in society and economy raises many intriguing policy and research questions. Transport is at the crossroads of strong but antagonistic interests, which can be summarised under the following headings: continued economic growth, fair geographical accessibility, and favourable quality of life. These research and policy issues in the transportation sector are certainly not unique for the STELLA-STAR partners but are also abundantly existent in many other countries, regions and cities in our world.

5. References

- AXHAUSEN, K. (2005). Social networks and travel: Some hypotheses. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 90-108). London: Ashgate.
- BAE, C. & Mayeres, I. (2003, January). *Equity considerations in transportation policies*. Paper presented at the second meeting of STELLA Focus Group 3, Arlington, Virginia.
- COUNCIL of the European Union. (2001). Draft joint report on social inclusion. *DG Employment and Social Affairs* (report 15223/01). Brussels: Author.
- DGTREN (European Commission Directorate-General for Energy and Transport, in cooperation with Eurostat). (2002). *European Union Energy and Transport in Figures 2001*. European Commission.
- DONAGHY, K., Poppelreuter, S. & Rudinger, G. (Eds.). (2005). *Social dimensions of sustainable transport: Transatlantic perspectives*. London: Ashgate.

- EUROPEAN Conference of Ministers of Transport. (2000). *Transport and ageing of the population. Report of the 112th Roundtable on Transport Economics* (Held in Paris on 19th–20th November 1998). Paris: OECD.
- GAFFRON, P. (2005). The implementation of walking and cycling policies. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp.195-208). London: Ashgate.
- HAKAMIES-Blomqvist, L. (1993). Compensation in older drivers reflected in their fatal accidents. *Accid Anal Prev*, 26, 107–112.
- HEAL, G. (1993). *Valuing the future: Economic theory and sustainability*. New York: Columbia University Press.
- INTERGOVERNMENTAL Panel on Climate Change. (2001). *Climate Change 2001: Impacts, Adaptation, and Vulnerability: Summary for Policymakers* [Online report]. Retrieved September 8, 2009, from http://www.grida.no/publications/other/ipcc_tar/?src=/climate/ipcc_tar/wg2/005.htm
- JONES, R.W. & Kierzkowski, H. (2001). A framework for fragmentation. In S.W. Arndt & H. Kierzkowski (Eds.), *Fragmentation: New production patterns in the world economy*. New York: Oxford University Press.
- KAPLAN, D. & Holloway, G. (1998). Segregation in Cities. In Association of American Geographers (Eds.), *Resource Publications in Geography*. Washington, DC: AAG.
- KRUGMAN, P. (2002, October 20). For Richer: How the permissive capitalism of the boom destroyed American equality. *The New York Times Magazine*, pp. 62-76.
- KWAN, M.-P. (2003, January). *New information technologies, human behavior in spaceTime, and the urban economy*. Paper presented at STELLA TRB Session 335, Washington, DC.
- LACKSHMANAN, T. R. (2001, March). *Sustainable transit with a focus on developing countries*. Paper presented at workshop on European and North American Perspectives on Sustainable Transportation, University of Illinois at Urbana-Champaign.
- MARCELLINI, F. & Allen, S. (2004, May). *Personal behaviour and sustainable transport: mobility problems and transportation needs of older people in Europe and the US*. Presented at the third meeting of STELLA Focus Group 3, Lisbon, Portugal.
- MEDDA, F. & Boarnet, M.G. (2005). An Analysis of the effects of urban land use on transportation. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp.127-143). London: Ashgate.
- MUNROE, D. K. & Hewings, G. J. D. (2007). The role of intraindustry trade in interregional trade in the Midwest of the U.S. In R. Cooper, K. Donaghy & G. Hewings (Eds.), *Globalization and Regional Economic Modeling* (pp.87-105). Berlin: Springer.

- PICKUP, L. & Giuliano, G. (2005). Transport and social exclusion in Europe and the USA. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 48-49). London: Ashgate.
- ROSENBLUM, S. & Stahl, A. (2002). *Automobility among the elderly: The convergence of environmental, safety, mobility and land use issues*. Presented at the meeting of STELLA Focus Group 4 Meeting, Helsinki, Finland.
- ROSENBLUM, S. & Hakamin-Blomqvist, L. (2004, May). *Women and sustainable transport*. Paper presented at the third meeting STELLA Focus Group 3, Lisbon, Portugal.
- RUDINGER, G. (2002). Mobility behavior of the elderly. In W.R. Black & P. Nijkamp (Eds.), *Social change and sustainable transport* (pp. 157-164). Bloomington: Indiana University Press.
- RUDINGER, G. & Jansen, E. (2003). Self-initiated compensations: results and suggestions from the study AEMEIS. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Ageing independently: Living arrangements and mobility* (pp. 220-233). Berlin, Heidelberg, New York: Springer.
- SCHAEIE, K.W. & Mollenkopf, H. (2005). Mobility issues in the United States and Europe. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 50-66). London: Ashgate.
- SCHINTLER, L. (2005). Society, behavior, and private/public transport: Trends and prospects in North America. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 29-37). London: Ashgate.
- SPECIAL Report: Demography and the West. (2002, August 24). *The Economist*, pp. 20-22.
- STERN, E. & Richardson, H. (2005). A new research agenda for modelling travel choice and behaviour. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp.144-163). London: Ashgate.
- SUCHORZEWSKI, W. (2005). Society, behavior, and private/public transport: Trends and prospects in transition economies of Central and Eastern Europe. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 14-28). London: Ashgate.
- TACKEN, M. (2004). Trip making behaviour among elderly in time and space. In H. Mollenkopf, F. Marcellini, I. Ruoppila & M. Tacken (Eds.), *Age and outdoor mobility. A European study* (pp. 81-101). Amsterdam: IOS.
- THE European Commission (2003). *Europe at a crossroads: The need for sustainable transport*. Brussels: Authors.
- VICKERMAN, R. (2001, March). *Transport intensity and the debate on decoupling transport growth from economic growth*. Paper presented at workshop on European and North American Perspectives on Sustainable Transportation, University of Illinois at Urbana-Champaign.

- WALDORF, B. & Pitfield, D. (2005). The effects of demographic shifts on non-automobile transportation needs of the elderly. In K. Donaghy, S. Poppelreuter & G. Rudinger (Eds.), *Social dimensions of sustainable transport: Transatlantic perspectives* (pp. 67-89). London: Ashgate.
- WILLIS, S.J. (2003). Programmatic interventions with the older driver: a conceptual framework. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Ageing independently: living arrangements and mobility* (pp. 263-278). Berlin, Heidelberg, New York: Springer.
- WORLD Business Council for Sustainable Development. (2001). *Mobility 2001. World mobility at the end of the twentieth century and its sustainability*. Retrieved September 8, 2009 from http://www.wbcds.org/web/projects/mobility/english_full_report.pdf
- WORLD Commission on Environment and Development (1987). *Our common future (The Brundtland Report)*. New York: Oxford University Press.

Georg Rudinger & Kristina Kocherscheid

Verkehrssicherheitsperspektiven für eine in die Jahre kommende Gesellschaft¹

Abstract

Mobil sein zählt zu den Grundbedürfnissen und im Rahmen des westlichen Wertesystems zu den Grundrechten von Menschen. Dies gilt umso mehr, je mobiler die Gesamtgesellschaft wird und je mehr Menschen im Zuge der demographischen Entwicklung zum Kreis der »älteren Menschen« zählen. Dementsprechend wertet der dritte Bericht zur Lage der älteren Generation in der Bundesrepublik Deutschland den Erhalt der Mobilität im Alter als Basis für die »Erschließbarkeit der verschiedensten Ressourcen der Außenwelt« und als »entscheidenden Faktor von Lebensqualität im Alter«, d.h. Aufrechterhaltung der *Unabhängigkeit der Lebensführung* durch Bewahrung von Fertigkeiten inklusive der Ermutigung zur Mobilität spielt für Ältere eine besonders wichtige Rolle.

Die Entwicklung von Interventionsansätzen zur Förderung einer lebenslangen, sicheren Mobilität älterer Menschen stellt eine gesamtgesellschaftliche Herausforderung dar. Dementsprechend lebendig ist die Diskussion über Möglichkeiten, ältere Verkehrsteilnehmer/innen in der Erhaltung ihrer Mobilität zu unterstützen und die Forschungsaktivität auf diesem Gebiet. Im Beitrag werden u.a. folgende Bereiche thematisiert:

- Senioren stellen keine homogene Gruppe dar; dieses gut erforschte Faktum gilt auch für ihr Mobilitätsverhalten, aber auch für das Risikoverhalten. Ältere Menschen sind durch ausgesprochene *Vielfalt mobilitätsbezogener Erwartungen, Verhaltensmuster und Lebensstile* charakterisiert.
- Die gegenwärtig durchaus wachsende positive Bewertung und *Einstellung gegenüber Alter und Altern* muß bei der Entwicklung neuer Sicherheitsmodelle berücksichtigt werden, gar Vorteil sollte daraus gezogen werden.
- Um höhere Straßenverkehrssicherheit – nicht nur für Ältere – zu erreichen, ist eine stärkere Orientierung auf schwache, ungeschützte, *vulnerable Verkehrsteilnehmer* notwendig.
- Die Teilnahme am Straßenverkehr ist Ergebnis eines lebenslangen Lernprozesses. *Mobilitätsgewohnheiten*, die während des Lebenslaufes erworben wurden, werden auch im Alter (unter Einsatz diverser *Kompensationsstrategien*) zum Großteil beibehalten.

¹ Vortrag auf dem ADAC/BAST-Symposium am 13. Oktober 2006 in Baden-Baden.

- Technologische Systeme (z.B. *Telematik*) können Mobilitätsplanung erleichtern und (bis zu einem gewissen Grad) Verluste in Fertigkeiten und Fähigkeiten zur Verkehrsteilnahme kompensieren und auf diese Weise zur Unfallreduktion beitragen. Die *Alterssensitivität* der Systeme muß ein zentraler Entwicklungsbereich sein.
- Es sollte eine sehr viel *stärkere Beteiligung der Älteren* an der Entwicklung, Einführung und Implementation von verkehrsbezogenen Sicherheitsmaßnahmen und neuen Technologien ermöglicht werden.
- Augenmerk sollte auch auf »alternative« Mobilitätsformen gelenkt werden: *Smart Modes* (zu Fuß gehen, Radfahren). Auch die Nutzung des ÖPNV sollte dadurch erleichtert werden, daß er so »benutzerfreundlich« wie nur möglich wird.
- Ältere haben weiterhin ein hohes *Interesse an* und eine positive Einstellung zu *Lernen, Übung und Training*. Dies gilt auch für technischen Fortschritt und moderne Entwicklungen bezüglich der Straßenverkehrssicherheit.
- Wenn Verkehrssicherheitsmaßnahmen implementiert werden, sollten die Kommunikationsmöglichkeiten dergestalt genutzt werden, daß sie für die verschiedenen Lebenssituationen und Lebensstile der Älteren angemessen sind. »Neue« Medien können dabei durchaus in Betracht gezogen werden. Fachdisziplinen wie Politikwissenschaft, Verkehrswesen, Ingenieurwissenschaften, Raumplanung, Architektur, Soziologie, Medizin, Psychologie, Pädagogik, Rechtswissenschaften, Ökonomie und Ökologie müssen künftig kooperieren, um einen Synergie-Effekt im Bereich der Mobilitätsförderung und der Verkehrssicherheitsarbeit für ältere Verkehrsteilnehmer/innen zu erzielen. Dabei ist eine enge, interdisziplinäre Zusammenarbeit notwendig – ebenso wie ein vorurteilsfreier Umgang der Menschen miteinander in unserer Gesellschaft.

1. Mobilität im Alter – ein Zukunftsthema

Die stetig zunehmende Zahl älterer Verkehrsteilnehmer/innen stellt die Verkehrssicherheitsarbeit vor immer neue Herausforderungen. Hintergrund ist nicht nur die allgemeine demographische Entwicklung, sondern auch eine breite Orientierung Älterer an einem aktiven Lebens- und Freizeitstil, in dessen Kontext die Mobilität eine wesentliche Rolle spielt (Gaffron, 2002; Kocherscheid & Rudinger, 2005). Die Möglichkeit, sich mit Hilfe auch eines Kraftfahrzeuges Ressourcen der Außenwelt zugänglich zu machen und Erlebnishorizonte zu erschließen, wird vielfach als zentrales Element von Lebensqualität im Alter betrachtet (BMFSFJ, 2001, S. 240).

1.1 Demographischer Wandel

»Die mobile Gesellschaft der Zukunft wird eine ergraute, internationale Gesellschaft sein« heißt es lapidar in einem rezenten Trend-Szenario (Ruff, 2003). Der demografische Wandel ist unabweisbar und für das nächste Dreivierteljahrhundert irreversibel (Birg, 2005). So wird sich in den nächsten 50 Jahren der Anteil älterer Menschen an der Gesamtbevölkerung weltweit verdoppeln (VCÖ, 1999). Durch das Altern der »Baby-Boom« Generation (zwischen 1946 und 1964 geboren), eine höhere Lebenserwartung sowie rückläufige Geburtenraten wird sich die Bevölkerungsstruktur in den nächsten 30 Jahren grundlegend verändern (»Explosion« bei den Älteren, »Implosion« bei den Jüngeren) (Birg, et al., 1998). Bis 2050 wird sich in den meisten OECD-Ländern der Anteil der über Achtzigjährigen voraussichtlich verdreifachen (OECD, 2002). Neben dem generellen Anstieg des Anteils älterer Menschen in der Bevölkerung nimmt derzeit mit jeder Alterskohorte auch die Zahl der Führerscheininhaber deutlich zu. In den nachrückenden Kohorten sind Führerscheinbesitz und eine lange Fahrpraxis weit verbreitet. Dieser Zuwachs wird durch Frauen allen Alters sowie vor allem durch ältere Menschen zustande kommen (Deutsche Shell GmbH, 2001). Während 1999 noch ca. 10 Prozent aller 80jährigen Frauen den Führerschein besaßen, wird sich dieser Anteil im Jahr 2025 voraussichtlich auf 80 Prozent erhöht haben (Pfefferott, 1994; Oswald, 1999). Bei den 60-64jährigen Frauen hat sich der Anteil in den letzten Jahren bereits von 11 Prozent auf 75 Prozent erhöht (vgl. Limbourg & Reiter, 2001). Die alternde Gesellschaft wird also zunehmend auch eine mobile Gesellschaft (auf Rädern) sein (vgl. Abbildung 1 und Abbildung 2).

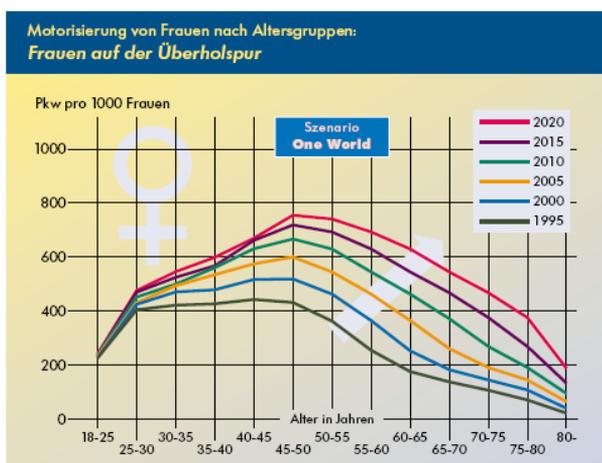


Abbildung 1: Motorisierung von Frauen nach Altersgruppen (Shell-Studie, 2001)

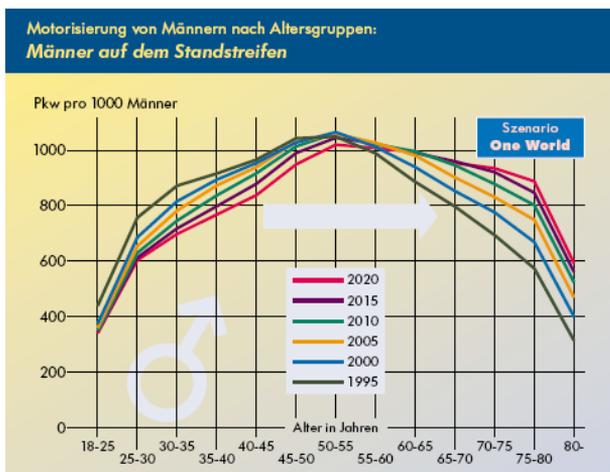


Abbildung 2: Motorisierung von Männern nach Altersgruppen (Shell-Studie, 2001)

Darüber hinaus ist abzusehen, dass immer mehr ältere Menschen auch in Zukunft einen aktiveren Lebens- und Freizeitstil zeigen bzw. ihren Lebensstil aus jüngeren Jahren fortschreiben werden, der bisher weitgehend nur mit jüngeren Altersgruppen assoziiert war. So wird Mobilität in all ihren Varietäten auch bei den Älteren zunehmend eine größere Rolle spielen (vgl. Bauer Verlagsgruppe, 1999). Dies führt dazu, dass sowohl die Automobilindustrie als auch die Anbieter öffentlicher Verkehrsmittel und nicht zuletzt auch Raum- und Verkehrsplaner vor neue Herausforderungen gestellt werden. Schließlich wird sich jeder einzelne Bürger mit der Frage zu beschäftigen haben, wie er seine persönliche Mobilität bei möglicherweise eingeschränkten physischen und psychischen Kompetenzen und Möglichkeiten im Alter gewährleisten möchte. Der Verkehrsexpertentag² fordert in diesem Zusammenhang, dass lebenslanges Informieren, Lernen und Reflektieren in Kooperation mit allen Beteiligten angestrebt werden müsse und erhofft sich dadurch folgende wünschenswerte Charakteristika einer angemessenen Mobilität in unserer Gesellschaft: Verbesserung der Beförderungs- und Transportqualität für Personen und Güter und der Standortqualität von Städten und Regionen, Erhöhung der Entscheidungs- und Planungssicherheit und der Effektivität von Mobilität und Verkehr, Erleichterung von Investitionsentscheidungen und von gezieltem Umweltschutz und nicht zuletzt Verbesserung der Lebensqualität. Durch diese Quadratur des Kreises soll zumindest unsystematischem Vorgehen und schädlichen Optimierungen einzelner Bestandteile von Mobilitäts- und Verkehrssystemen entgegenwirkt werden.

2 Deutscher Verkehrsexpertentag am 1. und 2. Juli 2004 in Bonn. *Mobilitäts- und Verkehrskultur*. [<http://www.verkehrsexpertentag.de>].

Dementsprechend ist eine gesamtgesellschaftliche Aufgabe, integrierte Ansätze zu entwickeln, um eine lebenslange, sichere Mobilität zu fördern und die künftigen Erfordernisse der alternden Bevölkerung so weit in das Bewusstsein der Öffentlichkeit zu rücken, dass sie anderen Prioritäten der Politik finanziell gleichgestellt werden (OECD, 2002). Dabei muss im Vordergrund stehen, dass durch die Förderung der lebenslangen Mobilität Kosteneinsparungen erzielt werden können (OECD, 2002). Zentral dabei ist, dass ältere Verkehrsteilnehmer/innen – zumindest teilweise – andere Mobilitätsbedürfnisse und –probleme haben als jüngere. Will man ältere Verkehrsteilnehmer/innen im Straßenverkehr vor den verkehrsbedingten Gefahren schützen, möglichst ohne ihre Mobilität einzuschränken, muss man ihre Mobilitätsbedürfnisse und –probleme kennen, um geeignete Verkehrssicherheitsmaßnahmen planen zu können. In diesem Feld haben die am Psychologischen Institut der Universität Bonn durchgeführten Studien AE-MEIS (Aeltere Menschen im Straßenverkehr), FRAME (Freizeitverkehr älterer Menschen), und VeBo (Verkehrssicherheitsbotschaften für Senioren) grundlegende Einsichten und Erkenntnisse gebracht.

1.2 Mobilität und Lebensqualität

Fragen von Mobilität und Lebensqualität stehen derzeit europaweit im Interesse der Forschung, dies spiegelt eine Vielzahl internationaler Studien wider. Hierzu zählen beispielsweise das von der EU 2003 eingerichtete Forschungsprojekt SIZE (<http://www.size-project.at> (Februar 2006)): »Life Quality of Senior Citizens in Relation to Mobility Conditions«, welches in acht europäischen Ländern die Mobilitätsbedingungen und Lebensqualität älterer Menschen untersucht, oder auch das Forschungsprojekt CARMA (<http://www.compass-org.at> (Februar 2006)): »Care for the Aged at Risk of Marginalization«, das als ein sozialwissenschaftliches Forschungsprojekt untersucht, wie ältere Menschen in einem unabhängigen Leben zu Hause unterstützt werden können, und das ebenfalls von der Europäischen Kommission im fünften Rahmenprogramm »Quality of Life« gefördert wird.

Mobilität ist somit zunehmend auch im höheren Lebensalter als ein Faktor anzusehen, der wesentlich für die Fähigkeiten jedes Einzelnen ist, ein selbstbestimmtes Leben auf möglichst hohem qualitativen Niveau zu führen.

Nach einer vom Bundesministerium für Familie, Frauen, Senioren und Jugend 2001 herausgegebenen Studie zur Erhaltung von Mobilität im Alter (Mollenkopf & Flaschenträger, 2001, S. 68ff.) lässt sich die Bedeutung von Mobilität im Alter in die folgenden Hauptaspekte gliedern:

- Mobilität als Bewegung an sich, als menschliches Grundbedürfnis,
- Mobilität, Fortbewegung als soziales Bedürfnis, als Wunsch nach gesellschaftlicher Integration und Teilhabe,

- die Möglichkeit der Fortbewegung als Ausdruck persönlicher Autonomie und Freiheit,
- Mobilität als Quelle neuer Eindrücke, als Anregung und Abwechslung,
- Mobilität als reflektierter Ausdruck von (noch vorhandener) Lebenskraft.

Möglichst uneingeschränkt mobil zu sein gehört demnach zu den wesentlichen Voraussetzungen für die Zufriedenheit älterer Menschen (Fookon, 1999; Mollenkopf & Flaschenträger, 2001; Kaiser, 2003). In diesem Zusammenhang kommt dem Erhalt der Mobilität und der sicheren Verkehrsteilnahme älterer Menschen ein großer Stellenwert zu (BMFSFJ, 2001). Dabei geht es nicht nur um Automobilität, sondern beispielsweise auch um die Nutzung öffentlicher Verkehrsmittel und sogenannte »smarte« Fortbewegungsarten wie Fahrradfahren oder zu Fuß gehen.

1.3 Kompetenzen und Einbußen älterer Verkehrsteilnehmer/innen

Die Entwicklung im Alter wird vielfach als ein heterogener und individuell gestaltbarer Prozess beschrieben (z.B. Baltes & Carstensen, 1996; Kaiser, 2003). Die älteren Kraftfahrer/innen bilden keineswegs eine einheitliche Gruppe. Auf diesen Umstand wird heute »eher plakativ« verwiesen (Kaiser, 2003), dennoch findet dieses Faktum erst zögerlich Eingang in das gesellschaftliche Bewusstsein. Bereits vor 20 Jahren wies Lehr auf diesen Umstand hin, erst in jüngster Zeit jedoch vollzieht sich eine Trendwende (Lehr, 2001).

Die Leistungsfähigkeit des Kraftfahrers kann als ein wesentlicher Aspekt im multikausalen Ursachengefüge bei der Verkehrssicherheit gesehen werden (Becker & Albrecht, 2003). So ist allgemein bekannt, dass der menschliche Alternsprozess verkehrsrelevante Gesundheits- und Leistungseinbußen mit sich bringt. Ältere Menschen leiden häufiger an chronischen Erkrankungen, die Bedeutung für die Verkehrsteilnahme haben, wie z.B. Herz-Kreislauf-Erkrankungen, Bluthochdruck, Diabetes, Wirbelsäulenleiden oder Demenzerkrankungen (Statistisches Bundesamt, 2000). Auch die oft sehr intensive Medikation mit Analgetika (Schmerzmittel), Sedativa (Beruhigungsmittel), Hypnotika (Schlafmittel) und Psychopharmaka (z.B. Antidepressiva) sowie die damit einhergehenden Wechsel- und Nebenwirkungen können die psychophysische Leistungsfähigkeit älterer Verkehrsteilnehmer/innen erheblich beeinträchtigen (Püllen, 2000). Aus der einschlägigen Literatur (vgl. Meyers et al., 1999; Owsley et al., 1999; De Raedt & Ponjaert-Kristoffersen, 2000; Kaiser & Oswald, 2000) ist ferner bekannt, dass etwa ab dem 75. Lebensjahr entweder altersbedingt oder in Verbindung mit einer Krankheit eine Reihe von Leistungseinbußen auftreten, die sich ungünstig auf die Verkehrssicherheit auswirken können. Hierzu zählen eine Erhöhung der Blendempfindlichkeit, eine abnehmende Sehschärfe in der Dämmerung, eine eingeschränkte Sichtweite bei Dunkelheit und eine Verschlechterung des Hörvermögens. Auch die Aufmerksamkeit, die Konzentrations- und Merkfähigkeit können

deutlich nachlassen, sowie die wichtige Fähigkeit, mehrere Aufgaben gleichzeitig durchzuführen. Sicherheitsrelevant ist auch eine Veränderung der Reaktionsgeschwindigkeit sowie eine Verengung des Sehbereichs (Holte & Albrecht, 2004), die dazu führt dass Objekte, die sich von links nähern, unter Umständen nicht mehr (rechtzeitig) wahrgenommen werden können (Kaiser & Oswald, 1999; Fozard, 2000; Jansen et. al, 2001). Zu diesen Beeinträchtigungen können sich Veränderungen der Persönlichkeit gesellen, wie z.B. zunehmende Ängstlichkeit, Rigidität, Nachlassen des Urteilsvermögens und der Bereitschaft zur Selbstkritik sowie eine Diskrepanz zwischen Fremd- und Selbstbild (Kaiser, 2000).

Es bleibt jedoch festzuhalten, dass Art und Ausmaß der psychophysiologischen und psychischen Veränderungen bei älteren Verkehrsteilnehmern große interindividuelle Variationen zeigen. Die Zugehörigkeit zu einem bestimmten chronologischen oder kalendarischen Alter besagt an für sich noch nichts, da sich Individuen gleichen chronologischen oder kalendarischen Alters auf der Ebene ihres biologisch-funktionalen Alters erheblich voneinander unterscheiden können, und das bedingt auch eine erhebliche Varianz in Bezug auf die Leistungsfähigkeit (Wahl et. al, 2000). Darüber hinaus können viele ältere Menschen den mit dem Alter potenziell zunehmenden Einschränkungen in verschiedenen Lebensbereichen mit zuträglichen kompensatorischen Strategien begegnen (Baltes et. al., 1996, 1998; Kruse, 1996; 1999; GEO Magazin, 2002, S. 1). Ob und wie sich verkehrsrelevante Veränderungen der Kompetenzen im Alter auf das Verhalten älterer Menschen im Straßenverkehr auswirken, soll im Folgenden aufgezeigt werden.

In der Verkehrsunfallstatistik schlägt sich ein erhöhtes Risiko älterer Personen bisher nicht nieder (vgl. Limbourg & Reiter, 2001). Langford, Methorst & Hakamies-Blomqvist (2006) verglichen in einer Studie das Unfallrisiko von Pkw-Fahrern unterschiedlicher Altersgruppen (18-20; 21-30; 31-64; 65-74; 75+). Die Befunde liefern Evidenz für den »Low mileage bias«, denn bei Berücksichtigung der jährlichen Fahrleistung zeigte sich unabhängig von der Altersgruppe bei niedriger Fahrleistung (unter 3000km/Jahr) eine sechsfach erhöhte Unfallrate im Vergleich zu Pkw-Fahrern mit einer hohen Fahrleistung (über 14000km/Jahr). Nach den Ergebnissen von Langford et al. (2006) nimmt die Unfallbeteiligung älterer Kraftfahrer/innen nur bei Fahrern zu, die weniger als 3000km/Jahr zurücklegen, und dieser Zuwachs ist erst ab dem 75. Lebensjahr beobachtbar. Ältere Pkw-Fahrer/innen sind somit, aus Sicht der Autoren, weder gefährdeter als andere Altersgruppen, noch stellen sie ein erhöhtes Risiko für andere Verkehrsteilnehmer/innen dar (vgl. Abbildung 3).

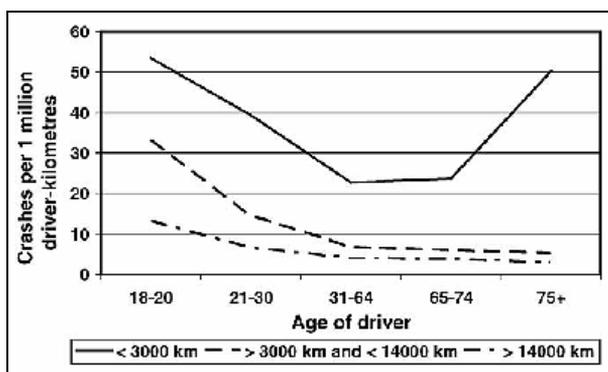


Abbildung 3: Jährliche Unfallbeteiligung unterschiedlicher Altersgruppen unter Berücksichtigung der jährlichen Fahrleistung (Langford, Methorst & Hakamies-Blomqvist, 2006, S. 576)

Sind ältere Menschen an einem Unfall beteiligt, so sind allerdings die Folgen besonders schwerwiegend. Berücksichtigt man die geringere Exposition Älterer als Pkw-Fahrer/innen, so sehen die relativen Unfallzahlen vor allem im höheren Alter (75+) nicht mehr so günstig aus. Dies weist darauf hin, dass insbesondere den 75jährigen und älteren Pkw-Fahrern zumindest im Mittel eine ausreichende Kompensation der im Alter einhergehenden Leistungsbeeinträchtigungen bei der Teilnahme am motorisierten Straßenverkehr immer schwerer fällt (Hakamies-Blomqvist et al., 1996). Zudem ist das Risiko zwischen den verschiedenen Arten der Verkehrsteilnahme deutlich unterschiedlich (Ernst, 1999; Kaiser, 2000). So sind beispielsweise von allen getöteten Fußgängern ca. 50 Prozent älter als 60 Jahre. Auch als Fahrradfahrer sind die Unfallfolgen für ältere Menschen vergleichsweise schwerwiegender als für andere Altersgruppen (Oswald, 1999).

Durch Verkehrsübertretungen fallen ältere Menschen derzeit weniger auf als jüngere. Auffällig sind ältere Pkw-Fahrer/innen hingegen bei Vorfahrtsdelikten (Kaiser & Oswald, 2000; Holte, 2004). Die häufigsten polizeilich erfassten Unfallursachen sind: Unzureichende Verkehrstüchtigkeit (ohne Alkohol), Vorfahrtsfehler, Fehler beim Abbiegen, Wenden, Rückwärtsfahren, beim Ein- und Ausparken sowie falsches Verhalten gegenüber Fußgänger/innen an Fußgängerüberwegen. Als Unfallursachen unbedeutend sind für diese Altersgruppe zum einen das Fahren unter Alkoholeinfluss und zum anderen überhöhte Geschwindigkeit (Emsbach & Friedel, 1999).

So ist es weniger ein bewusst riskantes Fahrverhalten, dass bei älteren Pkw-Fahrern Unfälle hervorruft, sondern eher Wahrnehmungs- bzw. Koordinationsfehler, die zu einer Gefährdung führen. Diese Interpretation entspricht der These von Schade (2000, S. 18), der von einer mit dem Alter reduzierten »Risikoannahme« ausgeht, während ein »unkompensierter Altersabbau«

Fehlleistungen im höheren Alter wahrscheinlicher werden lässt. Problematisch werden Wahrnehmungs- und Informationsmechanismen jedoch vorrangig dann, wenn eine Selbstüberschätzung vorliegt, der betroffene Fahrer sich demnach mehr zutraut als er leisten kann. In solchen Fällen besteht die Gefahr, sich in Situationen zu begeben, die man nicht ausreichend kontrollieren kann. Hier steigt die Wahrscheinlichkeit von Unfällen dramatisch an. Gerade in der fehlerhaften Selbsteinschätzung der eigenen Fähigkeiten und Defizite könnte bei einigen älteren Fahrern eine wichtige Risikoursache liegen (vgl. Engeln, 2003).

Sofern auf der einen Seite teilweise ungünstige psychophysische Veränderungen im Alter existieren, auf der anderen Seite jedoch keine besondere Auffälligkeit im Unfallgeschehen und bei der Verkehrsdelinquenz vorliegt, stellt sich die Frage nach den Kompensationsmöglichkeiten für ungünstige Entwicklungen im Alter.

1.4 Kompensationsstrategien älterer Verkehrsteilnehmer/innen

Auch wenn der Erhalt der Mobilität älterer Menschen, der als wesentlicher Garant für ihre Lebenszufriedenheit zu verstehen ist, zunehmend ins Zentrum des öffentlichen Interesses rückt, verläuft die Diskussion über ältere Verkehrsteilnehmer/innen selten wertfrei. Sie wird stets mit dem Unterton einer möglichen Gefährdung geführt, die potenziell von älteren Kraftfahrer/innen auszugehen scheint. Die finnische Verkehrswissenschaftlerin Hakamies-Blomqvist, die sich in den letzten Jahren am intensivsten in zahlreichen Forschungsarbeiten und Veröffentlichungen mit der Rolle der älteren Verkehrsteilnehmer/innen auseinandergesetzt hat – genannt seien der Bericht für den US Transportation Research Board »Transportation in an Ageing Society: A Decade of Experience« (2001) und ihre Beiträge in dem OECD-Bericht »Ageing and Transport« (2001) sowie im EU-Rahmen ihr Beitrag »Ageing Europe: The Challenges and Opportunities for Transport Safety« (ETSC, 2003) – geht davon aus, dass es immer noch die dominierende öffentliche Wahrnehmung ist, dass ältere Kraftfahrer/innen eine erhebliche Bedrohung für die Verkehrssicherheit darstellen. Diese Wahrnehmung beruht wahrscheinlich auf dem Wissen um altersbedingte Leistungseinbußen, die sowohl physische als auch kognitive Kompetenzen betreffen (vgl. Hakamies-Blomqvist et al., 1996; Hakamies-Blomqvist, 2003).

Zu den psychologischen Kompetenzen Älterer zählen neben ihrer sensorischen, motorischen und kognitiven Leistungsfähigkeit auch ihre Fähigkeiten, mit psychophysischen Leistungseinbußen ausgleichend umzugehen und so ihre Mobilität weitestgehend risikoarm zu gestalten (vgl. Schlag & Engeln, 2001). In der Gerontologie wird dieses Potenzial durch das Zusammenspiel der drei Komponenten Selektion, Optimierung und Kompensation anschaulich erklärt. Das Modell »Selektive Optimierung mit Kompensation« (vgl.

Baltes & Baltes, 1989) zielt darauf ab, die generelle Leistungsfähigkeit zu erhalten und zu verbessern, indem man Aufgaben, Anforderungen und Ziele umsichtig wählt, gut erhaltene Fähigkeiten sowie Ressourcen benutzt und diese oft übt, um so die Bereiche auszugleichen, die schwächer geworden sind.

Schlag und Engeln (2001) stellen in ihrem Aufsatz verschiedene Kompensationsstrategien vor. Selektion meint dabei vor allem die Auswahl oder die Veränderung von Zielen. Darüber hinaus beziehen sich Optimierungs- und Kompensationsprozesse auf die Mittel, mit denen Ziele erreicht werden (Baltes, 1999). Es liegt nahe, dass dieses Modell auch auf das Mobilitätsverhalten Älterer anwendbar ist. Nachlassende Informationsaufnahme und -verarbeitungsfähigkeit bewirken z.B., dass ältere Menschen Zeiten, Orte und Umstände ihrer Verkehrsteilnahme selektieren. Erfahrungen der Überbelastung im Verkehr bewegen Ältere u.a. dazu, ihre Verkehrsteilnahme einzuschränken. Andererseits können Handlungsmöglichkeiten durch die Übung bestehender Fähigkeiten und Fertigkeiten, z.B. im Rahmen von Verkehrsprogrammen für ältere Menschen als Fußgänger/innen, Radfahrer/innen und auch Pkw-Fahrer/innen, optimiert werden. Zudem kann beispielsweise die Nutzung des Pkws eine Strategie sein, um alterskorrelierte, individuelle motorische Fähigkeitsverluste zu kompensieren (Krämer, 2004).

Die Möglichkeiten zur Kompensation kann man auch in Anlehnung an das Modell des niederländischen Psychologen Michon (1988) sehen, welches die Verkehrsteilnahme als eine hierarchisch geordnete Verbindung von Teilaufgaben auf drei Ebenen beschreibt (vgl. Brouwer, 1994; Pfafferott, 1994):

- strategische Ebene
- taktische Ebene und
- operationale Ebene.

Auf der *strategischen Ebene* erfolgen Entscheidungen vor Antritt einer Fahrt (z.B. die der Fahrtroute oder des Fahrtzeitpunkts). Diese Entscheidungen können von älteren Fahrern, vor allem wenn sie sich im Ruhestand befinden und infolgedessen in ihrer Zeiteinteilung flexibler sind, ohne zeitlichen Druck getroffen werden. Auf *taktischer Ebene* steht das Streben nach einem konstant niedrigen Risiko während der Verkehrsteilnahme durch antizipatorische Fahrmanöver auf Basis der Kenntnis über sich selbst, das Fahrzeug und die Erwartungen hinsichtlich künftiger Verkehrssituationen im Vordergrund (z.B. Verlangsamung der Fahrgeschwindigkeit, wenn ein Verkehrszeichen einen Schulbereich ankündigt). Der Zeitdruck ist auf dieser Ebene schwach bis mäßig ausgeprägt. Auf *operationaler Ebene* geht es schließlich um die unmittelbare Auswahl und Ausführung von Manövern zur Gefahrenabwehr, etwa das Spurwechseln oder das plötzliche Auftauchen von Verkehrshindernissen. Es besteht hier ein konstanter Zeitdruck, da nur begrenzte Zeit für die Meidung oder Bewältigung gefährlicher Situationen zur Verfügung steht. Vor allem auf dieser Ebene sind Probleme bei den älteren

Pkw-Fahrern zu erwarten. Sie sind in diesen Situationen umso mehr überfordert, je mehr das sichere Verhalten eine selten praktizierte Fahrverhaltens-technik erfordert. Kompensationen älterer Kraftfahrer/innen sind daher in erster Linie auf taktischer und strategischer Ebene möglich, da ihnen dann meist ausreichend Zeit für Entscheidungen und Handlungen zur Verfügung steht.

Anhand dieses Modells lässt sich die relativ günstige Unfallsituation älterer Kraftfahrer/innen gut erklären. Ältere Kraftfahrer/innen versuchen Leistungseinbußen vor allem durch Vermeidung ungünstiger Verkehrszeiten und -situationen sowie durch eine defensive, vorsichtige und langsame Fahrweise auszugleichen (Schlag, 2003; Weinand, 1997).

Im »Outdoor-Mobility-Survey« gab beispielsweise jeder zweite ältere Autofahrer an, nicht zu Hauptverkehrszeiten zu fahren, wenn es nicht unbedingt erforderlich sei (Mollenkopf & Flaschenträger, 1997). Ein ähnliches Bild zeigt sich auch in den USA: »Older drivers reduce their nighttime driving, and, with increasing age, they also restrict their driving to lower speed roads..., low traffic volumes and good weather« (Waller, 1996, 2001).

Schon früh hat Böcher (1977; zit. nach Rühle, 1996) darauf hingewiesen, dass u.a. durch Erfahrung und Routine sowie durch entsprechende Merkmale der Gesamtpersönlichkeit (z.B. erhöhte Lernmotivation und Anstrengungsbereitschaft, Vorsichtshaltung, Selbstkritik, Einsicht in eigene Leistungsgrenzen) ein Ausgleich vorhandener Schwächen möglich ist.

Von verschiedenen Seiten ist auf den engen Zusammenhang zwischen Fahrerfahrung und Kompensation von Leistungsmängeln im höheren Lebensalter aufmerksam gemacht worden (vgl. Böcher, 1977; Seib, 1990). Zugute gehalten wird den Älteren als ausgleichender Faktor für altersspezifische Einschränkungen zumeist ihre vermehrte Erfahrung. Tatsächlich kann diese Erfahrung trotz eventuell eintretender Leistungsmängel genutzt werden, um die Risiken des Autofahrens möglichst gering zu halten.

Diese vermehrten Erfahrungen als Pkw-Fahrer/in müssten sich in einer verbesserten Fähigkeit ausdrücken, möglicherweise gefährliche Situationsentwicklungen frühzeitig richtig aufgrund nur geringer Hinweise vorauszusehen und ohne weiteres Zögern angemessen reagieren zu können. Im Verlauf der Praxis müssten dann die Verhaltensweisen in solchen bereits »erfahrenen« Situationen immer mehr optimiert werden, damit sie immer weniger variabel, immer exakter und immer schneller ausgeführt werden können. In einer komplexen Handlungsfolge wie dem Autofahren scheinen die Handlungen dann ineinander überzugehen, so dass sie nicht mehr im einzelnen bewusst gesteuert werden müssen, leichter, schneller und selbstverständlicher ablaufen und »automatisiert« werden. Solche positiven Wirkungen vermehrter Erfahrung sind bei älteren Pkw-Fahrern jedoch nur in Grenzen feststellbar. Vielmehr wünschen sich gerade ältere Menschen mehr Informationen, bevor sie eine Entscheidung fällen, und sie benötigen mehr Zeit zur Ent-

scheidungsfindung. Sie sind vorsichtiger bei der Handlungsvorbereitung – vielleicht als Kompensation für tatsächliche oder vermutete negative Entwicklungen bei der Wahrnehmung und Informationsverarbeitung. Sie wollen ihre Handlungen auch beim Autofahren bewusst überwachen, und genau dies verlangsamt die Aufmerksamkeitszuwendung zu neuen Situationsentwicklungen. Andererseits kann eine solche verstärkte Handlungskontrolle die Möglichkeiten zur Fehlerkorrektur verbessern, vorausgesetzt, dass genügend Zeit vorhanden ist (Schlag, 2001). Problematisch ist in diesem Zusammenhang lediglich eine Überschätzung des Wertes von Routine und Erfahrung als Garantien sicheren Fahrens. »Es ist indes nur eine Minderheit älterer Kraftfahrer/innen, die zu dieser Überschätzung neigt und sich damit selbst in eine Überforderungssituation bringt. Die fahreignungsrelevanten Veränderungen der Leistungsfähigkeit werden in der Regel von älteren Pkw-Fahrern bemerkt« (Chaloupka, 1994, S. 330).

Als weiterer wichtiger Ausgleichsfaktor wird in der Literatur die Persönlichkeit genannt. Wie aus zahlreichen wissenschaftlichen Untersuchungen und aus der medizinisch-psychologischen Begutachtungspraxis bekannt, ist Autofahren letztlich eine Sache des Charakters, der Persönlichkeit und des Verantwortungsbewusstseins. Die im Alter vorherrschende geringere Risikobereitschaft in Kombination mit einer gefestigten Persönlichkeit und mit einer niedrigen Tendenz zu aggressiven Interaktionen sowie zu emotionalem Autofahren kann erfolgreiche Kompensationsstrategien im Alter unterstützen (Weinand, 1997; Burgard, 2005).

Bezogen auf das Kompensationsverhalten älterer Pkw-Fahrer/innen konnten Rudinger & Jansen (2003) in ihrer Studie AEMEIS (Ältere Menschen im Straßenverkehr) folgende Einsichten und Erkenntnisse identifizieren: Die Mehrzahl der älteren Menschen vermied das Fahren, wenn sie sich körperlich nicht fit fühlten. Gesundheitliche Aspekte und das wahrgenommene Alter waren die häufigsten Gründe für die Aufgabe des aktiven Fahrens.

Einer der Schwerpunkte der Ergebnisse war die Erkenntnis, dass gezielte Verkehrssicherheitsmaßnahmen vor allem auf die 54- bis 64jährigen ausgerichtet sein sollten, da diese Altersgruppe sich einerseits durch eine relativ hohe Verkehrsbeteiligung und andererseits durch eine häufiger auftretende Tendenz auszeichnete, erkannte Leistungseinbußen zu ignorieren und ihr gewohntes – in diesem Fall unangemessenes – Fahr- und Mobilitätsverhalten beizubehalten (Gruppe der »dysfunktionalen Kompensierer«). Ältere Pkw-Fahrer/innen, bei denen bereits Leistungsdefizite bestanden, die ihr Fahrverhalten jedoch mit dem Ziel der Risikoreduktion umstellten, fielen in die Gruppe der »funktionalen Kompensierer«. Alle älteren Pkw-Fahrer/innen die ihr Fahrverhalten umstellten, obwohl keinerlei Defizite vorlagen, gehörten der Gruppe der »präventiven Kompensierer« an. Eine weitere Gruppe dieser Systematisierung wies keine Leistungsdefizite auf und zeigte (folglich) auch keine Änderungen im Fahrverhalten.

Insgesamt zeigten 23,5 Prozent der älteren Pkw-Fahrer/innen einen funktionalen Kompensationsstil, 5,5 Prozent einen dysfunktionalen und 25 Prozent einen präventiven Kompensationsstil. 30 Prozent zeigten ein stabiles Fahrverhalten und 16 Prozent der Befragten konnten nicht klassifiziert werden.

Ein Vergleich zwischen Personen mit funktionalem und dysfunktionalem Kompensationsstil konnte zeigen, dass beide Gruppen gleich häufig an Unfällen beteiligt waren, Personen mit funktionalem Stil jedoch wesentlich häufiger unverschuldet in die Unfälle verwickelt waren als dysfunktional kompensierende Fahrer/innen. In ihrem Einstellungsprofil unterschieden sich die beiden Gruppen dahingehend, dass die funktional kompensierenden Fahrer/innen mehr Einsicht in die Notwendigkeit hatten, ihr Fahrverhalten umzustellen, und dass sie weniger risikofreudig in ihrem Fahrverhalten sind. Auch zeigte sich eine Tendenz zur Selbstüberschätzung bei Personen mit dysfunktionalem Kompensationsstil. Die Frage nach dem Lebensstil der Kraftfahrer/innen konnte aufzeigen, dass dysfunktional kompensierende Personen eher einen Lebensstil führen, bei dem die Suche nach Aufregung eine wichtige Rolle spielt. Im Gegensatz dazu hatten die funktional kompensierenden Fahrer breite Interessen und einen moderaten Lebensstil.

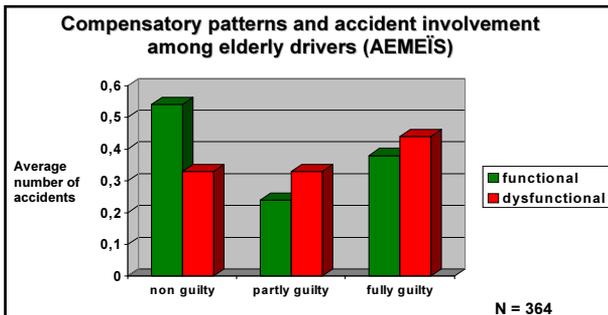


Abbildung 4: Kompensationsstrategien und Unfallbeteiligung älterer Kraftfahrer/innen

Des Weiteren war auffällig, dass besonders die Jüngeren unter den männlichen älteren Pkw-Fahrern (54 bis 64 Jahre) zu der Gruppe der dysfunktional kompensierenden Autofahrer gehörten. Deren Merkmale ähnelten sehr den Eigenschaften, die bei jungen, männlichen Pkw-Fahrern zu finden waren. Zusammenfassend kann gesagt werden, dass lediglich 5 bis 6 Prozent der älteren Pkw-Fahrer/innen, also diejenigen mit dysfunktionalem Kompensationsstil, keine Einsicht in eine Anpassung ihres Fahrstils an die eingeschränkte Fahrtüchtigkeit haben.

Inwieweit Kompensationsmöglichkeiten im konkreten Einzelfall vorhanden sind oder nicht, kann nur individuell unter Berücksichtigung der persönlichen Lebensumstände, Sichtweisen und vorliegenden Erfahrungen unter-

sucht werden (Höfner & Anderle, 1987). Auf Grund der Bedeutung von Mobilität im Alter sind Verhaltensveränderungen gegenüber Einschränkungen vorzuziehen. Allerdings darf neben reinen Verhaltensänderungen für eine risikoarme Mobilität im Alter jedoch die Rolle der Umweltverhältnisse nicht unberücksichtigt bleiben. Kompetenzen im Alter sind das Resultat der Wechselwirkungen von Personen- und Umweltmerkmalen. Voraussetzung für die Aufrechterhaltung der Verkehrsteilnahme älterer Menschen ist im Sinne des psychologischen Kompetenzmodells daher eine adäquate Gestaltung der Verkehrsumwelt. Merkmale der Verkehrsumwelt, wie z.B. der Einsatz neuer Techniken, können die individuelle Mobilität und Selbständigkeit älterer Menschen fördern. Für Fußgänger können z.B. Bordsteinabsenkungen, Fahrtrampen und Aufzüge die Überwindung von Höhenunterschieden erleichtern. Gut erreichbare Mittelstreifen und Verkehrsberuhigung können neben akustischen Ampelsignalen und langen Grünphasen die Fahrbahnüberquerung erleichtern. Niedrige Einstiegshöhen und ausreichende Einstiegs- und Umsteigezeiten könnten neben gut lesbaren Fahrplänen und leicht bedienbaren Fahrkartenautomaten die Nutzung des öffentlichen Personenverkehrs erleichtern (vgl. Krämer, 2004).

2. Interventionen im Bereich Mobilität und Verkehr

Grundlegende Ziele für Interventionen im Bereich Mobilität und Verkehr sind einerseits der »Abbau von Mobilitätsbarrieren« und andererseits der »Aufbau von Mobilitätskompetenz« (Kaiser, 1999, S. 263). Kaiser betont, dass Mobilität nicht nur die körperliche, sondern auch die geistige Dimension beinhaltet, und zwar die »geistige Beweglichkeit in kognitiven Möglichkeitsräumen«. Der Aufbau von Mobilitätskompetenz erfordert mehr als z.B. das Besorgen von technischen Hilfsmitteln, »denn Mobilität beginnt im Kopf« (Kaiser, 1999, S. 261).

Im Sinne einer größtmöglichen Effizienz von Maßnahmen zur Förderung mobilitätserhaltender Strategien erscheint ein mehrdimensionales Vorgehen am ehesten erfolgsversprechend. Unterschiedliche Maßnahmen sollten sich gegenseitig ergänzen. Sie sollten einerseits durch den älteren Menschen selbst und andererseits durch eine unterstützende Umwelt ergriffen werden. Für den älteren Menschen erscheint die Identifikation und Ausschöpfung der vorhandenen internen und externen Ressourcen zur Mobilitätserhaltung notwendig. Hierbei spielt die Eigeninitiative eine wichtige Rolle, die gleichzeitig einen großen Beitrag zur Erhaltung eines positiven Selbstbildes leistet. Die Optimierung vorhandener Fähigkeiten und Fertigkeiten kann zu einer wesentlichen Verbesserung der Mobilitätssituation älterer Menschen beitragen. Nicht zuletzt sollte eine kritische Selbsteinschätzung die Aktivitäten Älterer im Bereich der Mobilität begleiten. Mobilitätsunterstützende Aktivitäten der Umwelt sollten die Bereitstellung flexibler, bedarfsgerechter Mobilitätsange-

bote für ältere Menschen umfassen. Wichtig erscheint in diesem Zusammenhang eine umfassende Information der Älteren über bestehende Angebote im Mobilitätsbereich.

Die Erhaltung der Mobilität im Alter ist jedoch gegen das Risiko einer möglichen Verkehrsgefährdung im Einzelfall sorgfältig abzuwägen. Information und Beratung im Sinne von Verkehrssicherheitsarbeit mit Senioren sind notwendig, um diese in ihren Bemühungen um die Aufrechterhaltung der für das Führen eines Kraftfahrzeuges erforderlichen körperlichen und geistigen Gesundheit zu unterstützen. So kann die Bewahrung der mobilitätsbezogenen Sicherheit älterer Menschen einen Beitrag zum Erhalt der Mobilität älterer Menschen liefern (Kocherscheid et al. (im Druck)).

2.1 Kompensation durch technisch-konstruktive Gegebenheiten

Zur Optimierung vorhandener Ressourcen könnten ergonomische und/oder technische Unterstützungen für ältere Pkw-Fahrer/innen herangezogen werden, die geeignet sind, die Erfüllung alltäglicher Fahraufgaben zu erleichtern und ggf. vorhandene Leistungsminderungen auszugleichen. So kann man über eine konsequent durchdachte Gestaltung der Verkehrsumwelt vielen Leistungseinschränkungen begegnen. *»Ein konsequentes Umdenken wäre angebracht, das vom adaptiven Verhalten des Lenkers ausgeht, aber seinen Leistungsgrenzen durch die Konstruktion einer menschengerechten Umwelt Rechnung trägt«* (Cohen 2001, S. 241).

Färber (2000) gibt einen Überblick über die Unterstützung älterer Pkw-Fahrer/innen durch die technische Fahrerassistenz. Man kann z.B. ABS (Anti-Blockier-System), ESP (Elektronisches Stabilitätsprogramm), VES (Sichtverbesserungen im Straßenraum), elektrohydraulische Bremsen und andere ähnliche Systeme dazu zählen (Färber, 2003). Ein Schlüsselfaktor für den Einsatz dieser Systeme ist die Zuverlässigkeit der Funktion in allen Situationen und die Ausfallsicherheit.

Solche technologischen Systeme (z.B. telematische Sicherheitssysteme) können die Mobilitätsplanung erleichtern, Verluste in Fertigkeiten und Fähigkeiten älterer Verkehrsteilnehmer/innen kompensieren und auf diese Weise zur Unfallreduktion beitragen, sofern sie auf die Bedürfnisse und Einschränkungen des Alters abgestimmt sind (vgl. Rudinger, 2002). Die Alterssensitivität der Systeme muss demnach ein zentraler Entwicklungsbereich sein. Der Umgang mit diesen technologischen Systemen erfordert von den älteren Menschen jedoch ein gewisses Maß an kognitiven Ressourcen sowie Eigeninitiative und Interesse (Rudinger & Jansen, 2003). In diesem Zusammenhang kann man sich fragen, ob es generell möglich ist, durch die Vermittlung sinnvoller Kompensationstechniken oder durch die Förderung im Bereich kognitiver Basisprozesse die Fähigkeit Älterer im Umgang mit Technik zu steigern (Willis, 2003). Ein solcher Ansatz ist der sogenannten

personenorientierten Intervention zuzurechnen, in dem das Benutzermodell dem Designmodell angepasst wird. Die technikorientierte Intervention hingegen passt das Design- dem Benutzermodell an, demzufolge die Technik den Bedürfnissen des Nutzers entsprechend konstruiert wird (vgl. auch Charness & Bosman, 1995).

Des Weiteren sollten Forderungen an die Gestaltung neuer technologischer Systeme aufgestellt werden. Denkbar wäre eine seniorenerechte Gestaltung dieser Systeme (Gerontotechnik), die altersbedingte Einschränkungen berücksichtigt. Alternativ wäre jedoch auch ein »transgenerational design« (vgl. Rudinger & Jansen, 2003) denkbar, welches für alle Altersgruppen gleichermaßen von Nutzen wäre.

In Anbetracht der technischen Möglichkeiten dieser Systeme stellt sich die Frage, welche Auswirkungen sie tatsächlich auf die Verkehrssicherheit insbesondere älterer Verkehrsteilnehmer/innen haben. Denn häufig genug kann es gerade die Technik sein, die Barrieren schafft, Abhängigkeiten aufbaut und Gefühle der Hilflosigkeit hervorruft. Insgesamt sollte also berücksichtigt werden, dass technische Erleichterungen durchaus zu einem »trägerischen Sicherheitsgefühl« führen können, was ein riskanteres Fahrverhalten zur Folge haben kann. Inwieweit dies jedoch für ältere Pkw-Fahrer/innen zutrifft, wurde noch nicht überprüft (vgl. Färber, 2003).

Auch auf die Schaffung einer altersgerechten Verkehrsumwelt wird an vielen Stellen hingewiesen (z.B. Kaiser & Oswald, 1999; Schneider & Holz-Rau, 2001). Im Vordergrund stehen hier vor allem die »Entschleunigung« des Verkehrs durch bauliche oder verkehrsregulierende Maßnahmen und die übersichtliche Verkehrsführung.

2.2 Möglichkeiten der Defizitkompensation durch Training und Beratung

Weitere Vorschläge zur Ausschöpfung vorhandener Fähigkeiten und Fertigkeiten wenden sich an Pkw-Fahrer/innen selbst. Sie stützen sich auf die Möglichkeit, durch Training eventuellen Fähigkeitsverlusten entgegenzutreten. Dabei muss der Tatsache Rechnung getragen werden, dass man trotz möglicher Leistungseinschränkungen bei älteren Menschen nicht von einem generellen irreversiblen Abbau der Kräfte im Alter sprechen kann. Menschen altern in verschiedenen Funktionsbereichen unterschiedlich und viele dieser Prozesse lassen sich durch gezieltes Training verlangsamen, aufhalten oder sind für das sichere Führen eines Kraftfahrzeuges irrelevant. Prinzipiell kann man nachweisen, dass physische, psychische und soziale Funktionen weitgehend erhalten bleiben, solange man sie gebraucht: »Use it or lose it!«. Vor diesem Hintergrund kann man durchaus sagen: Ältere Fahrer trainieren die relevanten intellektuell-kognitiven, psychophysischen und psychomotorischen Funktionen bereits, indem sie weiter Auto fahren. Es gilt, dass ein permanentes Training in Form regelmäßiger, nach Möglichkeit tagtäglicher

Verkehrsteilnahme als Kraftfahrer/in das Optimum in Bezug auf Erhaltung und Förderung der Funktionsbereitschaft darstellt, sofern nicht krankheitsbedingte, schwere Leistungseinbrüche erfolgen.

Übung und Training wird in der Literatur auch im Rahmen des Disuse-modells diskutiert (vgl. z.B. Kausler, 1991), wobei dieses Alternsmodell sich sowohl auf physische als auch auf psychische und soziale Funktionen bezieht. Altern wird dann nicht als schicksalhaft ablaufender Abbauprozess verstanden, sondern kann durch Training und Gebrauch der körperlichen, psychischen und sozialen Funktionen beeinflusst werden (Olbrich, 1987). Baltes & Baltes (1989) gehen in ihrem psychologischen Modell davon aus, dass sich die Leistungsreserven mit zunehmendem Alter verringern. Durch die Strategie der Optimierung durch Selektion (von subjektiv bedeutsamen Bereichen) und durch Kompensation (von funktionalen Defiziten) gelingt es jedoch den älteren Menschen, trotz reduzierter Leistungsressourcen, die Aufgaben in den für sie relevanten Lebensbereichen zu bewältigen.

Im Rahmen von verkehrspädagogischen Programmen³ für ältere Menschen, sogenannte Moderatorenprogramme, wird in angeleiteten Gesprächskreisen versucht, die Probleme des Autofahrens im Alter aufzuarbeiten und die eigene, individuelle Situation zu reflektieren. Ähnliche Programme gibt es in Finnland (Schrey, 2000) und in Dänemark (Lund, 2000). Als besonders kritisch kann jedoch angesehen werden, dass solche Programme nicht generationsübergreifend wirken, dass die Programme den Schwerpunkt nicht auf die Motivation der Senioren legen und dass die Programme nicht verhaltensorientiert arbeiten (Becker et al., 2001). Zugleich muss festgestellt werden, dass der Wirkungsgrad der Programme nicht befriedigend ausfällt, da vor allem jene angesprochen werden, die bereits ein bestimmtes Problembewusstsein erreicht haben (vgl. Kaiser et al., 1998). Zusammenfassend bleibt also anzumerken, dass die Wirkung der Breitenaufklärung von Senioren zum Thema automobile Verkehrssicherheit durch die existierenden Seminarprogramme größtenteils noch unbefriedigend verläuft. Gerade vor dem Hintergrund, dass die Zielgruppen meist sehr groß gefasst werden (ab 55, 60 oder 65 Jahren) entstehen aufgrund der Heterogenität der Seminarteilnehmer/innen Probleme und möglicherweise Interessenskonflikte. Dies rückt den Ansatz der personenzentrierten Verkehrsaufklärung/-beratung älterer Kraftfahrer/innen in den Vordergrund, da hier im Gegensatz zu solchen Programmen, bei denen die Moderatoren nur eine Momentaufnahme der Befindlichkeit und Leistungsfähigkeit der Teilnehmer/innen erhalten, eine kontinuierliche Beobachtung und

3 Der Deutsche Verkehrssicherheitsrat (DVR) bietet durch sog. Umsetzerverbände wie die Deutsche Verkehrswacht oder den ADAC das Programm »Ältere aktive Kraftfahrer (ÄaK)« und »Ältere Menschen als Fußgänger« (ÄMaFIS) an, dessen Durchführung vom Bundesministerium für Verkehr, Bau- und Wohnungswesen finanziert wird.

eine abgestimmte Beratung aufgrund der im Alter vorliegenden schleichenden Veränderungen und starken Schwankungen der Tagesform erstrebenswerter erscheint. Des Weiteren sollte auch das private Umfeld der zu Beratenden berücksichtigt werden, da es Aufschluss über die Relevanz der eigenen Mobilität und über mögliche Fortbewegungsalternativen zum Auto gibt, so dass Rückschlüsse auf die Motivation der informierten Senioren gezogen werden können. Auch kann bei einer kontinuierlicheren Beratung und Begleitung der Wirkungsgrad der Beratung und mögliche resultierende Einstellungs- und Verhaltensänderungen überprüft werden. Für eine solche kontinuierliche Beratung, in der über eine Momentaufnahme hinausgehende Informationen über den Gesundheitszustand und das private Umfeld älterer Menschen berücksichtigt werden können, kommen vor allem Hausärzte als professionelle Ansprechpartner/innen infrage, die in der Regel das Vertrauen der Senioren haben. Der Hausarzt trägt des Weiteren zur Identifikation von Risikopersonen beitragen, denn er ist hinsichtlich physischer und psychischer Veränderungsprozesse, die die Mobilitätstauglichkeit beeinflussen, kompetent. Im Rahmen einer gezielten Individualaufklärung sollten ältere Verkehrsteilnehmer/innen dementsprechend über altersspezifische Leistungsmöglichkeiten, die Auswirkungen von Medikamenten auf die Verkehrstüchtigkeit, Unfallrisiken und Möglichkeiten zur Kompensation aufgeklärt werden. Zudem sollen die älteren Menschen lernen, ihr Fahrverhalten den eigenen Möglichkeiten anzupassen. Allerdings müssten Ärzte auf eine solche Funktion im Rahmen ihres Behandlungsalltages gezielter als bisher vorbereitet und ausgebildet werden. Ziel des Forschungsprojektes VeBo (Verkehrssicherheitsbotschaften für Senioren) ist daher die Entwicklung einer ärztlichen Fortbildungsmaßnahme, die der Steigerung der verkehrsmedizinischen Kompetenz behandelnder Ärzte dienen soll. Es geht hierbei gezielt um Hilfestellung für ältere Patienten und um die Thematisierung der mit ihrer Erkrankung assoziierten Problemfelder. Durch die Erhöhung der kommunikativen Kompetenz des Arztes/der Ärztin soll zudem die Vermittlung entsprechender Diagnosen und ärztlicher Empfehlungen erleichtert und für den Patienten eher annehmbar gemacht werden. Die Einbindung von Vertrauenspersonen (Angehörigen) soll eine unterstützende Funktion bei der Unterrichtung von älteren Patienten erfüllen mit dem Ziel, eine sichere Verkehrsteilnahme Älterer zu gewährleisten, und somit einen deutlichen Beitrag zur Erhöhung der Verkehrssicherheit von Senioren leisten (Kocherscheid et al., (im Druck)).

3. Schlussfolgerung

Nicht nur unter Gerontologen ist unstrittig, dass Mobilität ein bedeutsamer Aspekt der Lebensqualität im Alter und sogar eine deren Voraussetzungen ist. Allerdings ist die Mobilität im Alter durch verschiedene Faktoren bedroht: Der Verlust an körperlicher Beweglichkeit, die Verschlechterung psy-

chischer und psychomotorischer Funktionen, technische, bauliche und soziale Bedingungen erzeugen Hemmnisse und Barrieren. Der für ältere Menschen so notwendige Erhalt der Mobilität erfordert es, diese Hemmnisse so weit wie möglich zu minimieren oder ganz zu beseitigen, wobei die Maßnahmen zum Wohle älterer Menschen sich selbstverständlich an der Art des Problems sowie der Zielgruppe orientieren müssen. Eine der möglichen Problemlösungen besteht in der Förderung automobiler Mobilität auch für Menschen im höheren Lebensalter. Dabei sollten Kampagnen zur Förderung der Verkehrssicherheit älterer Menschen zielgruppenspezifisch ausgerichtet sein und eine individuelle Ansprache (z.B. Arzt-Patient-Interaktion im Rahmen einer Mobilitätsberatung) ermöglichen. Gleichzeitig sollten sie sich an empirisch ermittelten Mobilitätsmotiven/-bedürfnissen sowie der spezifischen Lebenssituation Älterer orientieren und dabei die Bedeutung von Mobilität für ältere Verkehrsteilnehmer/innen berücksichtigen. Die Forschung zeigt, dass eine solche Förderung keineswegs fahrlässig, sondern verantwortbar ist. Psychologen, Mediziner und Pädagogen, Ingenieure und Techniker, Städteplaner und Politiker finden hier ein zukunftsweisendes Aufgabenfeld. Zukunftsweisend ist es, weil sich nicht nur der Anteil älterer Menschen in unserer Gesellschaft überhaupt erhöht, sondern besonders und überproportional der Anteil derjenigen Älteren, die als Pkw-Fahrer/innen unterwegs sind. Notwendig ist in dieser Situation eine enge, fachübergreifende Zusammenarbeit aller einschlägigen Fachleute unter Mitarbeit der Betroffenen, nötig ist aber auch ein liberaler, vorurteilsfreier Umgang der Menschen miteinander in unserer Gesellschaft (Informationsdienst Psychologie, 2005). Ebenso notwendig sind anwendungsorientierte Forschungsprojekte und Transferaktivitäten, die diejenigen unterstützen, die ganz praktisch an der Lösung von Verkehrsproblemen arbeiten und sich um die Einführung von Innovationen im Verkehr bemühen. Nicht zuletzt entscheidet die soziale Einbettung von neuen Verkehrsangeboten einschließlich ihrer technischen Komponenten über den Erfolg. Es reicht nicht, dass Verkehrskonzepte gut gemeint und nach Ansicht ihrer Entwickler nützlich und technisch brillant sind. Sie bedürfen vor allem der Akzeptanz durch die Nutzerinnen und Nutzer, sie müssen in die Lebenswelten und Bedürfnislagen individualisierter Gesellschaften passen. *»Sozialwissenschaftliche Mobilitätsforschung tut Not, um Mobilität überhaupt zu verstehen und künftigen Verkehr besser bewältigen zu können«* (Frankfurter Rundschau Online, 2003).

4. Implikationen für die Praxis

Auf der Basis u.a. von Expertenworkshops (z.B. Rudinger et al., 1999; Becker et al., 2001), einer Konferenz zur Sicherheit von Senioren im Straßenverkehr (Köln, 2000), der Studie AEMEIS (Jansen, 2001; Rudinger & Jansen, 2003), eines Beitrages von Kroj (2003), der sich auf diese vorgenann-

ten Materialien stützt, aber auch unter Einbezug eigener neuerer Studien wie FRAME (Rudinger et al., 2004), STELLA (Donaghy et al., 2004) und VeBo (Kocherscheid et al. (im Druck)) lassen sich schon einige Zielvorstellungen zu Verkehrssicherheitsperspektiven vor allem Ältere betreffend formulieren:

Gesellschaftliche Rahmenbedingungen

- Die Aufrechterhaltung der Unabhängigkeit der Lebensführung durch die Bewahrung von Fertigkeiten inklusive der Ermutigung zur Mobilität spielt für Ältere eine besonders wichtige Rolle.
- Senioren stellen keine homogene Gruppe dar; dieses gut erforschte Faktum gilt auch für ihr Mobilitätsverhalten, aber auch für das Risikoverhalten. Ältere Menschen sind durch ausgesprochene Vielfalt mobilitätsbezogener Erwartungen, Verhaltensmuster und Lebensstile charakterisiert.
- Die gegenwärtig durchaus wachsende positive Bewertung und Einstellung gegenüber Alter und Altern muss bei der Entwicklung neuer Sicherheitsmodelle berücksichtigt werden.
- Über Medienkampagnen muss Informationsarbeit geleistet und für mehr öffentliches Bewusstsein gesorgt werden. Gleichzeitig lassen sich dadurch Vorurteile, die von älteren Verkehrsteilnehmern ein verzerrtes Bild vermitteln, ausräumen.
- Altersabhängige Pflichtüberprüfungen für ältere Kraftfahrer/innen scheinen am beabsichtigten Zweck vorbeizugehen. Empfohlen werden vielmehr Beratungsmaßnahmen, die auf Risikogruppen abzielen, da dadurch Einzelfälle gründlicher geprüft und Maßnahmen individuell zugeschnitten werden können.
- Zur Verbesserung des spezialpräventiven Zugriffs (z.B. Schaffung eines tragfähigen Ausgleichs zwischen den Mobilitätswünschen älterer Menschen einerseits und möglichen von ihnen ausgehenden Unfallrisiken andererseits) bietet es sich an, funktionelle Assessments als Orientierung für die Beurteilung der Fahrkompetenz älterer Verkehrsteilnehmer/innen zu entwickeln.

Verkehrsstrukturelle Intervention

- Um höhere Straßenverkehrssicherheit – nicht nur für Ältere – zu erreichen, ist eine stärkere Orientierung auf schwache, ungeschützte, vulnerable Verkehrsteilnehmer notwendig.
- Die Teilnahme am Straßenverkehr ist Ergebnis eines lebenslangen Lernprozesses. Mobilitätsgewohnheiten, die während des Lebenslaufes erworben wurden, werden auch im Alter (unter Einsatz diverser Kompensationsstrategien) zum Großteil beibehalten. Das heißt, dass Mobilität für Ältere mit Hilfe technischer Möglichkeiten und durch Gelegenheiten zu

Training entsprechender Fertigkeiten aufrechterhalten und gar gefördert werden sollte.

- Technologische Systeme (z.B. Telematik) können Mobilitätsplanung erleichtern und (bis zu einem gewissen Grad) Verluste in Fertigkeiten und Fähigkeiten zur Verkehrsteilnahme kompensieren und auf diese Weise zur Unfallreduktion beitragen. Die Alterssensitivität der Systeme muss ein zentraler Entwicklungsbereich sein.
- Augenmerk sollte auch auf »alternative« Mobilitätsformen gelenkt werden wie z.B. Smart Modes (zu Fuß gehen, Radfahren). Auch die Nutzung des ÖPNV sollte dadurch erleichtert werden, dass er so »benutzerfreundlich« wie nur möglich wird.

Personale Ressourcen

- Die Älteren sind ihre eigenen Experten bezüglich ihres Alltagslebens inklusive ihrer Alltagsmobilität. Es muss deshalb eine sehr viel stärkere Beteiligung der Älteren an der Entwicklung, Einführung und Implementation von verkehrsbezogenen Sicherheitsmaßnahmen und neuen Technologien ermöglicht werden.
- Ältere haben weiterhin ein hohes Interesse an und eine positive Einstellung zu Lernen, Übung und Training. Dies gilt auch für technischen Fortschritt und moderne Entwicklungen bezüglich der Straßenverkehrssicherheit.
- Wenn Verkehrssicherheitsmaßnahmen implementiert werden, sollten verschiedene Kommunikationsmöglichkeiten dergestalt genutzt werden, dass sie für die verschiedenen Lebenssituationen und Lebensstile der Älteren angemessen sind. »Neue« Medien können dabei durchaus in Betracht gezogen werden.

Literatur

- BALTES, M. & Carstensen, L. (1996). Gutes Leben im Alter. Überlegungen zu einem prozessorientierten Metamodell erfolgreichen Alterns. *Psychologische Rundschau*, 47, 199-215.
- BALTES, M., Lang, F. & Wilms, H. (1998). Selektive Optimierung mit Kompensation. Erfolgreiches Altern in der Alltagsgestaltung. In A. Kruse (Hrsg.), *Psychosoziale Gerontologie* (Bd. 1: Grundlagen); (*Jahrbuch der medizinischen Psychologie* (Bd. 15). Göttingen: Hogrefe.
- BALTES, M., Maas, I., Wilms, H. & Borchelt, M. (1996). Alltagskompetenz im Alter: Theoretische Überlegungen und empirische Befunde. In K.U. Mayer & P.B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S. 525-542). Berlin: Akademie Verlag.
- BALTES, P.B. (1999). Altern als unvollendete Architektur der Humanontogenese. *Zeitschrift für Gerontologie und Geriatrie*, 32, 433-448.

- BALTES, P.B. & Baltes, M. (1989). Optimierung durch Selektion und Kompensation. *Zeitschrift für Pädagogik*, 35, 85-105.
- BALTES, P.B. (2002). *Erfolgreiches Altern*. In GEO Magazin Nr. 08/02: Lebenslauf-Forschung.
- BAUER Verlagsgruppe und Axel Springer Verlag AG (Hrsg.). (1999). *Forever young? Die Verschiebung der Altersschwellen. Eine Trendanalyse aus der Verbraucheranalyse 99*. Hamburg.
- BECKER, S. & Albrecht, M. (2003). Verkehrsmedizinische Aspekte im Alter. *Zeitschrift für Gerontopsychologie und -psychiatrie*, 16 (3), 101-115.
- BECKER, S., Berger, R. & Dumbs, M. et al. (2001). Projektgruppe »Perspektiven der Verkehrssicherheitsarbeit mit Senioren«. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Berichte der Bundesanstalt für Straßenwesen: Mensch und Verkehr* (Heft M 131). Bremerhaven: Wirtschaftsverlag NW.
- BIRG, H. (2005) *Der lange Bremsweg. Grundkurs Demographie – letzte Lektion*. FAZ, 53, 4.03.2005, S. 37. Zugriff 10.09.2009 unter <http://www.faz.net/s/Rub117C535CDF414415BB243B181B8B60AE/Doc~EF504AE75F2264929B36F54C78975DE3D~ATpl~Ecommon~Scontent.html>
- BIRG, H., Flöthmann, E.J. & Frein, T. et. al. (1998). *Simulationsrechnungen zur Bevölkerungsentwicklung in den alten und neuen Bundesländern im 21. Jahrhundert*. Bielefeld.
- BÖCHER, W. (1977). Die verkehrspsychologische Beurteilung älterer Menschen. *Zeitschrift für Gerontologie*, 10 (3), 169-173.
- BROUWER, W. (1994). Ältere Autofahrer und Anforderungen an die Aufmerksamkeit. In U. Tränkle (Hrsg.), *Autofahren im Alter. Mensch-Fahrzeug-Umwelt* (Bd. 30, S.121-137). Köln/Bonn: TÜV Rheinland/Deutscher Psychologen Verlag.
- BUNDESANSTALT für Straßenwesen [Federal Highway Research Institute]. (2000). Deutscher Verkehrssicherheitsrat e.V., Bundesministerium für Verkehr, Bau- und Wohnungswesen, Kommission der Europäischen Union. Internationale Konferenz: Mehr Verkehrssicherheit für Senioren. 2000: 2.-4.05, Köln. Report M 123.] Reports of the Federal Highway Research Institute] Bergisch Gladbach, Germany: BASt.
- BUNDESMINISTERIUM für Familie, Senioren, Frauen und Jugend BMFSFJ. (2001). *Dritter Bericht zur Lage der älteren Generation in Deutschland*. Berlin: Autor.
- BURGARD, E. (2005). *Fahrkompetenz im Alter: Die Aussagekraft diagnostischer Instrumente bei Senioren und neurologischen Patienten*. Unveröffentlichte Dissertation, Ludwig-Maximilians-Universität zu München.
- CANZLER, W., Götz, K. et al. (2003). *Das Auto im richtigen Leben*. Frankfurter Rundschau Online. Zugriff am 10.09.2009 unter www.isoe.de/ftp/kog_schu_FR_10_03
- CHALOUPIKA, C. (1994). Was hat es mit der Veränderung der Leistungsfähigkeit älterer Verkehrsteilnehmer wirklich auf sich? In U. Tränkle (Hrsg.), *Autofahren*

- im Alter Mensch-Fahrzeug-Umwelt* (S. 321-334). Köln/Bonn: TÜV Rheinland/Deutscher Psychologen Verlag.
- CHARNESS, N., & Bosman, E.A. (1995). Compensation through environmental modification. In P.A. Dixon & L. Bäckman (Eds.), *Compensating for psychological deficits and declines* (pp 147-168). Mahwah, NJ: Lawrence Erlbaum.
- COHEN, A. (2001). Leistungsanforderungen und Möglichkeiten der Senioren als Fahrzeuglenker. In A. Flade, M. Limbourg & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (S. 241-258). Opladen: Leske und Budrich.
- DE RAEDT, R. & Ponjaert-Kristoffersen, I. (2000). The relationship between cognitive/neuropsychological factors and car driving performance in older adult. *Journal of the American Geriatric Society*, 48, 1664-1668.
- DEUTSCHE Shell GmbH. (2001). *Shell Pkw-Szenarien. Mehr Autos – weniger Verkehr? Szenarien des Pkw-Bestands und der Neuzulassungen in Deutschland bis zum Jahr 2020*. Hamburg.
- DONAGHY, K.P., Poppelreuter, S. & Rudinger, G. (Eds.). (2004). *Social Dimensions of Sustainable Transport: Transatlantic Perspectives*. London: Ashgate.
- EBERHARD, J., Hakamies-Blomqvist, L. & Howard, E. et al. (2001). *Ageing and transport: mobility needs and safety issues. Expert group on mobility needs and safety problems of an ageing society* (ERS 4). Paris: OECD.
- EMSBACH, M. & Friedel, B. (1999). Unfälle älterer Kraftfahrer. *Zeitschrift für Gerontologie und Geriatrie*, 32, 318-325.
- ENGELN, A. (2003). Zur Bedeutung von Aktivität und Mobilität für die Entwicklung im Alter. *Zeitschrift für Gerontopsychologie & -psychiatrie*, 16 (3), 117-129.
- EUROPEAN Transport Safety Council (ETSC). (2003). *Transport safety performance in the EU: A statistical overview*. Brussels: Authors.
- ERNST, R. (1999). *Mobilitätsverhalten und Verkehrsteilnahme älterer Menschen*. (Europäische Hochschulschriften, Reihe XXII, 340). Frankfurt a.M.: Peter Lang.
- FÄRBER, B. (2000). Neue Fahrzeugtechnologien zur Unterstützung der Mobilität Älterer. *Zeitschrift für Gerontologie und Geriatrie*, 33 (3), 178-185.
- FÄRBER, B. (2003). Microinterventions: Assistive devices, telematics, and person-environment interactions. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Aging independently; Living arrangements and mobility* (pp. 248-263). Berlin: Springer.
- FLEISCHMANN, U.M. (1991). Gedächtnis. In W.D. Oswald, W.M. Herrmann, S. Kanowski, U. Lehr & H. Thomae (Hrsg.), *Gerontologie* (S.168-176). Stuttgart: Kohlhammer.
- FOOKEN, I. (1999). Entwicklungsgegebenheiten außerhalb der Wohnung über die Lebensspanne. In H.W. Wahl et. al. (Hrsg.), *Konferenz »Alte Menschen in ihrer Wohnung«*. (Veranstalter Deutsches Zentrum für Altersforschung Wüstenrot Stiftung, S. 145-151). Opladen; Wiesbaden; Westdeutscher Verlag.

- FOZARD, J.L. (2000). Sensory and cognitive changes with age. In K.W. Schaie & M. Pietrucha (Eds.), *Mobility und transportation in the elderly* (pp. 1-61). New York: Springer.
- GAFFRON, P. (2002). The implementation of walking and cycling policies. *Paper presented at the meeting of STELLA Focus Group 3*, Bonn.
- HAKAMIES-Blomqvist, L. (2003). *Ageing Europe: The challenges and opportunities for transport safety*. Brussels: European Transport Safety Council (ETSC).
- HAKAMIES-Blomqvist, L. (2001). Safety of older persons in traffic. Transportation in an ageing society: A decade of experience. *Technical Papers and Reports from a Conference*. Washington, D.C.: Transportation Research Board.
- HAKAMIES-Blomqvist, L., Johansson, K. & Lundberg, C. (1996). Medical screening of older drivers as a traffic safety measure — A comparative Finnish-Swedish evaluation study. *Journal of the American Geriatrics Society*. 44, 650-653.
- HÖFNER, K.J. & Anderle, F.G. (1987). Probleme älterer Menschen bei der Teilnahme am Straßenverkehr. In G. Kroj & E. Spoerer (Hrsg.), *Wege der Verkehrspsychologie. Faktor Mensch im Verkehr* (Heft 36, S. 63-78). Braunschweig: RGG-Verlag.
- HOLTE, H. (2004). Automobil und sicher im Seniorenalter. *Zeitschrift für Verkehrssicherheit*, 1. 3-12.
- HOLTE, H. & Albrecht, M. (2004). Verkehrsteilnahme und -erleben im Straßenverkehr bei Krankheit und Medikamenteneinnahme. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit* (Heft M 162). Bremerhaven: Wirtschaftsverlag NW.
- INFORMATIONSDIENST Psychologie. (2005). Mobilität im Alter – ein Zukunftsthema. *IDP 3/2005*.
- JANSEN, E. (2001). Ältere Menschen im künftigen Sicherheitssystem Straße/Fahrzeug/Mensch [Elderly people in a future traffic system]. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit* (Heft M 134). Bremerhaven: Wirtschaftsverlag NW.
- KAISER, H.J. (Hrsg.). (1999). *Altern und Autofahren*. Bern: Huber Verlag.
- KAISER, H.J. (2000). Mobilität und Verkehr. In H.W. Wahl & C. Tesch-Römer (Hrsg.), *Angewandte Gerontologie in Schlüsselbegriffen*, 261-268. Stuttgart, Berlin, Köln: Kohlhammer.
- KAISER, H.J. (2003). Differentialpsychologisch-gerontologische Anmerkungen zur Verkehrssicherheitsarbeit mit Älteren. *Zeitschrift für Gerontopsychologie und -psychiatrie*, 16 (2), 131-140.
- KAISER, H.J. & Myllmäki-Neuhoff, J. (1998). Die Verkehrsteilnahme älterer Menschen als komplexes Handlungsproblem. Psychologische und pädagogische Grundfragen von Verkehrssicherheitsmaßnahmen für ältere Menschen. *Deutsche Verkehrswacht Forum* (Heft 1).

- KAISER, H.J. & Oswald, W.D. (2000). Autofahren im Alter – eine Literaturanalyse. *Zeitschrift für Gerontopsychologie und -psychiatrie*, 13 (3/4), 131-170.
- KAUSLER, D.H. (1991). *Experimental psychology, cognition and human aging*. New York: Springer.
- KRÄMER, K. (2004). Alt & Mobil: Kompetenzen älterer Verkehrsteilnehmer. *Presse-seminar »Senioren im Straßenverkehr«*, Leipzig 16.-17. September. Berlin: Deutsches Zentrum für Altersfragen.
- KOCHERSCHIED, K., Rietz, C. & Poppelreuter, S. (im Druck). Verkehrssicherheitsbotschaften für Senioren: Nutzung der Kommunikationspotenziale im allgemeinmedizinischen Behandlungsalltag. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit*. Bremerhaven. Wirtschaftsverlag. NW.
- KOCHERSCHIED, K. & Rudinger, G. (2005). Ressourcen älterer Verkehrsteilnehmer und Verkehrsteilnehmerinnen, In W. Echterhoff (Hrsg.), *Strategien zur Sicherung der Mobilität älterer Menschen* (S. 19-42). (Schriftenreihe Mobilität und Alter der Eugen-Otto-Butz-Stiftung). Köln: TÜV Verlag.
- KROJ, G. (2003). Future developments of mobility in the elderly: A German traffic psychologist's perspective. In K.W. Schaie, H. W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Aging independently: Living arrangements and mobility* (pp. 311-319). New York: Springer Publishing Company.
- KRUSE, A. (1996). *Kompetenz im Alter in ihren Bezügen zur objektiven und subjektiven Lebenssituation*. Darmstadt: Steinkopff.
- KRUSE, A. & Lehr, U. (1999). Reife Leistung. Psychologische Aspekte des Alterns. In A. Niederfranke, G. Naegele & E. Frahm (Hrsg.), *Funkkolleg Altern* (Bd. I, S. 187-238). Wiesbaden: Westdeutscher Verlag.
- LANGFORD, J., Methorst, R. & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash-risk – A replication of low mileage bias. *Accident Analysis and Prevention*, 38, 574-578.
- LEHR, U. (2000) Senioren – eine heterogene Gruppe. Demographische und psychologische Aspekte des Alterns. In Meyer-Hentschel Management-Consulting (Hrsg.), *Handbuch Seniorenmarketing*, 21-27. Frankfurt.
- LIMBOURG, M. & Reiter, K. (2001). Das Verkehrsunfallgeschehen im höheren Lebensalter. In A. Flade, M. Limbourg & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (pp. 211-227). Opladen: Leske und Budrich.
- LUND, W. (2000). Verkehrskampagnen und Kurse für ältere Fahrer in Dänemark. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit* (Heft M 123, S.102-103). Bremerhaven: Wirtschaftsverlag NW.
- MEYERS, J.E., Volbrecht, M. & Kaster-Bungaard, J. (1999). Driving is more than pedal pushing. *Applied Neuropsychology*, 6 (3), 154-164.
- MOLLENKOPF, H. & Flaschenträger, P. (2001). *Erhaltung der Mobilität im Alter*. Schriftenreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend (Band 197). Stuttgart: Kohlhammer.

- MOLLENKOPF H. & Flaschenträger, P. (1997) Mobilität im Alter. In H. Reents (Hrsg.), *Handbuch der Gerontotechnik* (12/97, Kap. III-4.6.2.). Landsberg: ecomed.
- OECD (2002). *Aging and transport: mobility needs and safety issues*. Paris: Author.
- OECD (2001). *Ageing and transport: mobility needs and safety issues*. Paris: Author.
- OLBRICH, E. (1987). Kompetenz im Alter. *Zeitschrift für Gerontologie*, 20, 319-330.
- OSWALD, W.D. (1999) Ältere Autofahrer. Täter oder Opfer? In H.J. Kaiser & W.D. Oswald (Hrsg.), *Altern und Autofahren*. Bern, Göttingen, Toronto, Seattle: Huber.
- OWSLEY, C., Ball, K. & Sloane, M. E. (1991). Visual cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging*, 6, 403-415.
- PFAFFEROTT, I. (1994). Mobilitätsbedürfnisse und Unfallverwicklung älterer Autofahrer/innen. In U. Tränkle (Hrsg.), *Autofahren im Alter* (S. 19-59). Köln: TÜV Rheinland.
- PÜLLEN, R. (2000). Autofahren. In I. Füsgen (Hrsg.), *Der ältere Patient – problemorientierte Diagnostik und Therapie* (S. 542-546). München: Urban und Fischer.
- RUDINGER, G. & Jansen, E. (2003). Technik, Neue Medien und Verkehr. In S.H. Filipp & U. Staudinger (Hrsg.), *Entwicklungspsychologie des mittleren und höheren Erwachsenenalters, Enzyklopädie der Psychologie, Themenbereich C Theorie und Forschung* (Serie V Bd.5, S. 559-597). Göttingen: Hogrefe.
- RUDINGER, G., Holz-Rau, C. & Grotz, R. (2004). Freizeitmobilität älterer Menschen. In Institut für Raumplanung (IRPUD) (Hrsg.), *Dortmunder Beiträge zur Raumplanung – Reihe Verkehr*. Dortmund.
- RUDINGER, G. & Jansen, E. (2003). Self-initiated compensations: Results and suggestions from the study AEMEIS. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Ageing independently: Living arrangements and mobility* (pp. 220-233). New York: Springer.
- RUDINGER, G. (2002). Mobility behavior of the elderly. In W.R. Black & P. Nijkamp (Eds.), *Social change and sustainable transport* (pp.157-164). Bloomington and Indiana: Indiana University Press.
- RUDINGER, G., Jansen, E., Jung, C. et. al. (1999, Mai). *Seniorenworkshop des Psychologischen Institutes der Universität Bonn*. Expertentagung im Auftrag der Bundesanstalt für Straßenwesen, Bonn.
- RUFF, F. (2003, Oktober). *Zukunft und Mobilität – Perspektiven der mobilen Gesellschaft. Journalisten und Wissenschaftler im Gespräch – Chemie für eine nachhaltige Mobilität am 06. und 07. Oktober 2003 in Münster*. (Presse-Information, Daimler Chrysler AG).
- RÜHLE, R. (1996). Alternde Menschen als Verkehrsteilnehmer. *Zeitschrift für Verkehrssicherheit*, 42, 52-61.
- SCHLAG, B. (2001). Ältere Menschen im Pkw unterwegs. In A. Flade, M. Limbourg & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (S. 85-98). Opladen: Leske und Budrich.
- SCHLAG, B. (2003). Safety and accidents among older drivers: The German perspective. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Ageing*

- independently: Living arrangements and mobility* (pp. 205-220). New York: Springer Publishing Company.
- SCHLAG, B. & Engeln, A. (2001). Kompensationsmöglichkeiten und Bewältigungsstrategien im Alter. In A. Flade, M. Limbourg & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (S. 259-273). Opladen: Leske und Budrich.
- SEIB, H. (1990). Erkenntnisse der Unfallursachenforschung zur altersbedingten Leistungsminderung der älteren Verkehrsteilnehmer, ihr Unfallrisiko, und die rechtlichen Konsequenzen. *Zeitschrift für Gerontologie*, 23, 86-96.
- SCHREY, A. (2000). Ältere Fahrer als Zielgruppe für die Verkehrssicherheitsarbeit in Finnland. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Schriftenreihe der Bundesanstalt für Straßenwesen (BASt): Mensch und Sicherheit* (Heft M 123, S.92-94). Bremerhaven: Wirtschaftsverlag NW.
- STATISTISCHES Bundesamt (Hrsg.). (2000). Diagnosedaten der Krankenhauspatienten 1998. *Fachserie 12* (Reihe 6.2). Wiesbaden: Statistisches Bundesamt.
- VERKEHRSLUB Österreich (VCÖ). (1999). *Senioren und Mobilität*. Wien: VCÖ.
- WAHL, H.W., Tesch-Römer, C. & Rott, C. (2000). Vision and cognitive functioning in old age. In B. Silverstone, M.A. Lang, B. Rosenthal & E. Faye (Eds.), *The Lighthouse Handbook of Vision Impairment and Vision Rehabilitation* (pp. 431-439). New York: Oxford University Press.
- WALLER, P.F. (2001). Future developments of mobility in the elderly – the American perspective. *Vortrag anlässlich der Conference »Aging in the Community: Living Arrangements and Mobility« 2001: 4.-6. April*. Heidelberg.
- WALLER, P.F. (1996). Accidents: Traffic. In J.E. Birren (Ed.), *Encyclopedia of gerontology. age aging and the aged* (1, pp. 19-25). San Diego: Academic Press.
- WEINAND, M. (1997). Kompensationsmöglichkeiten bei älteren Kraftfahrern mit Leistungsdefiziten. *Berichte der Bundesanstalt für Straßenwesen: Mensch und Sicherheit* (Heft M 77). Bremerhaven: BASt.
- WILLIS, S.J. (2003). Programmatic interventions with the older driver: A conceptual framework. In K.W. Schaie, H.W. Wahl, H. Mollenkopf & F. Oswald (Eds.), *Ageing independently: Living arrangements and mobility* (pp. 263-301). New York: Springer.

Kristina Kocherscheid

Konzeption und Evaluation einer ärztlichen Fortbildung zur Mobilitätsberatung älterer Kraftfahrer

Vor dem Hintergrund, dass in der Bundesrepublik Deutschland die Fahrerlaubnis nicht altersbegrenzt wird, erhält die gestiegene Mobilität älterer Menschen und ihre Beteiligung am Unfallgeschehen eine besondere Bedeutung für eine zielgruppenspezifische Verkehrssicherheitsarbeit. Die Berücksichtigung typischer Leistungseinschränkungen und Leistungseinbußen älterer Verkehrsteilnehmer spielt dabei unabhängig von der Art ihrer Verkehrsteilnahme eine wichtige Rolle. Es stellt sich die Frage, wie eine möglichst sichere Verkehrsteilnahme bei gleichzeitig zu erwartenden, individuellen, verkehrssicherheitsrelevanten Leistungsverminderungen in der heterogenen Gruppe der Senioren gewährleistet werden kann. Entsprechend stellen Beratungsangebote, die die Individualität der Senioren berücksichtigen, eine sinnvolle Alternative dar. In diesem Zusammenhang kann dem Hausarzt als wichtigem Ansprechpartner und kompetenter Vertrauensperson eine Schlüsselrolle zukommen. Er kann insbesondere zur Identifikation von Risikopersonen beitragen, für die individuell zugeschnittene Maßnahmen zur Mobilitätsgestaltung entwickelt werden müssen. Ziel muss hierbei immer das akzeptable Gleichgewicht zwischen Mobilität und Sicherheit (individuell wie generell) sein.

1. Ausgangslage

Die Sicherung der Mobilität älterer Menschen hat insbesondere in einer alternden Gesellschaft hohe Priorität. Mobilität ist die notwendige Voraussetzung vieler Aktivitäten, die der Sicherung der Lebensqualität und der sozialen Teilhabe dienen. Für den älteren Menschen ist Mobilität eine zentrale Voraussetzung dafür, ein selbstständiges Leben zu führen (Gaffron, 2002). Die Erhaltung und Förderung von Mobilität ist somit ein hochrangiges individuelles und politisches Ziel. Die Verbesserung der Verkehrssicherheit älterer Menschen stellt dabei ein Teilziel dar (Kocherscheid & Rudinger, 2005; Müller, 2003).

Der Mobilität älterer Menschen als Schlüssel zur Unabhängigkeit und zur Nutzung der persönlichen und umweltspezifischen Potenziale stehen jedoch – zumindest heutzutage noch – die mit höherem Lebensalter verbundenen verkehrsrelevanten Gesundheits- und Leistungseinbußen entgegen (Ewert, 2008). Mit zunehmendem Alter ist auch mit einem Abbau derjenigen Funkti-

onen (Seh- und Hörfähigkeit, kognitiver Funktionsbereich, Psychomotorik) zu rechnen, die für das rasche Erkennen von Geschehnissen im Straßenverkehr und für das angemessene Reagieren (unter Zeitdruck) vor allem auf komplexe Situationen erforderlich sind. Zu diesen Beeinträchtigungen können Veränderungen der Persönlichkeit hinzukommen (z. B. zunehmende Ängstlichkeit, Rigidität, Nachlassen des Urteilsvermögens und der Bereitschaft zur Selbstkritik) (Fozard, 2000; Wahl, Tesch-Römer & Rott, 2000).

Dennoch muss die Verkehrsteilnahme Älterer jenseits dieser einzugestehenden Defizite differenziert betrachtet werden, da es sich bezüglich der Fahreignung sowie verkehrsbezogener Einstellungen, Verhaltensweisen und Gewohnheiten um eine ausgesprochen heterogene Gruppe handelt. Altern ist eben kein einheitlich interindividuell invarianter Prozess. Die spezielle mobilitätsbezogene Sicherheit älterer Menschen ist ebenso wie die generelle Erhaltung ihrer Mobilität auch an individuelle Faktoren geknüpft, die sich zum Beispiel im persönlichen Lebensstil oder der Lebenslage widerspiegeln können. Entsprechend fällt das Gefährdungspotenzial im Straßenverkehr in der Gruppe der Älteren sehr unterschiedlich aus. Darüber hinaus können viele ältere Menschen Einschränkungen in verschiedenen Lebensbereichen mit kompensatorischen Strategien begegnen (Engeln & Schlag, 2008). Den Risikofaktor »Alter im Straßenverkehr« betreffend müsste somit nachgewiesen werden, dass das Kollektiv der älteren Kraftfahrer ein größeres Gefahrenpotenzial darstellt als andere Altersgruppen. In der aktuellen Verkehrsunfallstatistik schlägt sich ein erhöhtes Risiko älterer Kraftfahrer (bis 75 Jahre) bisher jedoch nicht nieder (Limbourg & Reiter, 2001).

Ältere fahren also nicht schlechter als Jüngere, sie fahren anders. Bemerkenswert ist auch, dass offensichtlich viele ältere Kraftfahrer in realistischer Einschätzung der eigenen Fahreignung das Autofahren von sich aus begrenzen oder ganz darauf verzichten (Kocherscheid, Rietz, Poppelreuter et al., 2007). Unter diesen Gesichtspunkten ist also eine Ungleichbehandlung, nach der nur für ältere Verkehrsteilnehmer die Fahrerlaubnis zeitlich begrenzt oder eine Verlängerung der Fahrerlaubnis von einer ärztlichen Untersuchung abhängig gemacht werden soll, nicht gerechtfertigt (Fellhase, 2005).

Es bleibt festzuhalten, dass die Durchführung einer altersbedingten gesetzlichen Überprüfung der Fahreignung nur dann gerechtfertigt wäre, wenn von dieser Altersgruppe eine erhebliche Gefahr ausgehen würde. Forderungen nach ärztlicher Untersuchung der Fahreignung aller Verkehrsteilnehmer über 60 oder 70 Jahre sind somit undifferenziert und nicht haltbar. Die Frage nach der Geeignetheit oder Ungeeignetheit älterer Verkehrsteilnehmer ist daher differenziert zu beantworten. Richtigerweise können gerade bei älteren Menschen – sofern sie nicht durch Unfälle auffällig wurden – selbst bei festgestellter Leistungseinschränkung Führerscheine unter Auflagen und Beschränkungen aufrechterhalten werden (Deutscher Verkehrsgerichtstag, 2009).

Hinzu kommt, dass Ärzte ohne geriatrisches und verkehrsmedizinisches Fachwissen kaum in der Lage sein werden, Risikofälle zu erkennen. So bleibt zu fragen, ob die mit der Einführung altersabhängiger Sonderbestimmungen im Fahrerlaubnisrecht verbundenen Kosten in einem vertretbaren Verhältnis zu der Verkehrsgefährdung stehen, die von älteren Verkehrsteilnehmern tatsächlich ausgeht, oder ob Zeit und Geld nicht vernünftiger in Maßnahmen zu investieren sind, die den älteren Menschen den Erhalt der für ihre selbstständige Lebensführung so wichtigen Mobilität erleichtern (Fellhase, 2005).

Vor diesem Hintergrund scheint es erstrebenswert, den bei einem Teil dieser Fahrergruppe auftretenden Leistungsbeeinträchtigungen und den daraus resultierenden Risiken durch individuelle Beratungsmaßnahmen ohne behördlichen Druck zu begegnen. Eine solche Vorgehensweise wird der Heterogenität der Zielgruppe gerecht, Einzelfälle können gründlicher geprüft und individuell zugeschnittene Maßnahmen zur Mobilitäterhaltung entwickelt werden, die sich an den spezifischen Lebenssituationen älterer Menschen orientieren und dabei die Bedeutung von Mobilität für den Einzelnen berücksichtigen. Multiplikator für solche Beratungsmaßnahmen sollten Ärzte sein, da es sich bei Einschränkungen der Fahreignung im Alter primär um medizinische Gründe handelt. Physische und psychische Veränderungsprozesse, die die Fahreignung beeinflussen, können am ehesten vom Hausarzt kompetent und objektiv beurteilt werden. Zudem wird sein Urteil von den Patienten respektiert und akzeptiert (Cooper & Snosna, 1983; Fastenmeier, Gsalter & Eggerdinger, 2005). Hervorzuheben ist dabei, dass der Arzt seine älteren Patienten in einem individuellen Gespräch direkt auf ihre durch physische und psychische Einschränkungen bedingte Gefährdung im Straßenverkehr ansprechen kann und durch die regelmäßig stattfindenden Konsultationen auch schleichende Veränderungen bemerkt. Somit hat er die Möglichkeit, die Umsetzung präventiver Verhaltensmaßnahmen unter konkreten Lebensbedingungen zu fördern und kann die Akzeptanz und Erfolge beim Patienten gut einschätzen (Starfield, 1998; Wippenbeck, 2002). Vor diesem Hintergrund liegt es nahe, eine Beratung in Fällen, in denen die Fahreignung älterer Kraftfahrer angezweifelt wird, in die Hände von (Haus-)Ärzten zu legen (Becker, Berger & Dumbs, 2001).

Dies setzt grundlegend voraus, dass der behandelnde Arzt neben geriatrischem und verkehrsmedizinischem Fachwissen, insbesondere hinsichtlich verkehrsrelevanter Erkrankungen, auch Kenntnisse über die rechtlichen Rahmenbedingungen der Verkehrsteilnahme älterer Menschen und der ärztlichen Aufklärungspflicht hat. In diesem Kontext ist es wichtig, dass ein behandelnder Arzt Erkrankungen, die eine Gefährdung für die Verkehrssicherheit darstellen können, in den komplexen Rahmen von medizinischer Hilfe, weiteren bestehenden Einschränkungen und Kompensationsmöglichkeiten einordnen kann. Zudem wird die Qualität der Arzt-Patient-Beziehung entscheidend durch die kommunikative Kompetenz des Arztes bestimmt. In der

Behandlung speziell älterer Menschen stellt das ärztliche Gespräch ein entscheidendes diagnostisches und therapeutisches Instrument dar (Innovations Report, 2001). Von Ärzten muss somit erwartet werden, dass sie, neben der nötigen Sachkompetenz für die Einschätzung der Mobilitätsprobleme und möglicher Bewältigungsstrategien, auch soziale Kompetenzen für die Gestaltung der Kommunikation mitbringen (Breder, 2001; Kalwitzki, 1996; Kocherscheid, 2008). Da eine Ausbildung spezieller kommunikativer Fertigkeiten für das Arzt-Patienten-Gespräch im Allgemeinen und den Umgang mit schwierigen Patienten im Besonderen nicht Teil des Medizinstudiums ist, ist jedoch anzuzweifeln, dass jeder Arzt über eine entsprechende Kompetenz verfügt. Zudem besteht ein erhebliches Defizit in der verkehrsmedizinischen Aus- und Weiterbildung der Ärzte (Hoffmann-Born & Peitz, 2006). Daher müssen Ärzte auf eine Funktion als »Lotse für ältere Verkehrsteilnehmer« im Rahmen ihres Behandlungsalltages gezielter als bisher vorbereitet und entsprechend ausgebildet werden, so dass sie ihre älteren Patienten angemessen verkehrsmedizinisch beraten können. Gerade vor dem Hintergrund der großen Relevanz kommunikativer Fähigkeiten ist eine explizite Schulung einer ausschließlich theoretischen Auseinandersetzung wie beispielsweise durch Handbücher (siehe hierzu Henning, 2007) vorzuziehen. Vor diesem Hintergrund stellt die Entwicklung eines ärztlichen Fortbildungskonzeptes zur Mobilitätsberatung älterer Patienten einen wesentlichen Eckpfeiler für die den Arzt zukünftig beschäftigenden Aufgaben im Umgang mit seinen älteren Patienten dar und bildet die wesentliche Basis für die Gestaltung der Beziehung zwischen alterndem Patient und Arzt (Kocherscheid et al., 2007).

2. Vorgehensweise

In diesem Kontext wurde im Rahmen des Forschungsvorhabens VeBo (Verkehrssicherheitsbotschaften für Senioren – Nutzung der Kommunikationspotenziale im allgemeinmedizinischen Behandlungsalltag) ein Fortbildungskonzept für Ärzte konzipiert, das die Ärzte unter Berücksichtigung der Bedingungen des Behandlungsalltages und der Bedürfnisse der zu behandelnden Senioren darauf vorbereiten soll, im Rahmen einer Mobilitätsberatung gezielt Aufklärungsarbeit hinsichtlich medizinisch relevanter Aspekte einer sicheren Verkehrsteilnahme zu leisten. Die behandelnden Ärzte sollen auf der Grundlage der in der Fortbildung vermittelten Inhalte Kernkompetenzen zu Fragen bezüglich »Verkehrssicherheit und Alter« erlangen, verkehrssicherheitsrelevante Informationen an ihre älteren Patienten weitergeben, sicherheitsrelevante Einstellungen zum Verkehrsverhalten fördern und somit einen deutlichen Beitrag zur Erhöhung der Verkehrssicherheit von Senioren leisten. Ein solches Vorgehen ermöglicht die frühzeitige Einleitung von Kompensationsmaßnahmen, um die Mobilität im Alter für den Kraftfahrer zu verlängern.

Die Grundlage für die Konzeption dieser Fortbildung bildete eine im Großraum Köln/Bonn durchgeführte Befragung an insgesamt knapp hundert Ärzten, Akteuren des Gesundheitswesens, Senioren und deren Angehörigen. Die vor dem Hintergrund dieser Ergebnisse konzipierte Fortbildung wurde vor ihrer Implementierung bezüglich inhaltlicher, konzeptueller und akzeptanzbezogener Aspekte evaluiert, um eine maximale Praxisorientierung und Nachhaltigkeit der im Rahmen der Fortbildung vermittelten Inhalte für den ärztlichen Berufsalltag zu gewährleisten. Die Ergebnisse dieser Evaluation legten eine interdisziplinäre Betrachtungs- und Vorgehensweise zur Verkehrsteilnahme älterer Menschen im Rahmen einer solchen Fortbildungsveranstaltung nahe, die neben medizinisch-psychologischen Grundlagen sowohl rechtliche Rahmenbedingungen der Verkehrsteilnahme älterer Menschen als auch kommunikative Aspekte der Arzt-Patient-Interaktion berücksichtigt. Sie unterstützen in weiten Teilen die Hoffnung auf eine erfolgreiche Implementierung der Fortbildung in den hausärztlichen Behandlungsalltag. Von allen teilnehmenden Ärzten wurden sowohl die Inhalte als auch die Organisation der Veranstaltung überaus positiv bewertet. Dieses Ergebnis spiegelt sich auch in der Globalbeurteilung der Fortbildungsveranstaltung wider, die von 64 Prozent der Teilnehmer mit »sehr gut« und von 36 Prozent mit »gut« bewertet wurde. Die Evaluationsergebnisse weisen zudem darauf hin, dass die Mediziner altersbedingten Einschränkungen der Fahreignung aufgrund physischer oder psychischer Leistungsminderung oder auch als Begleitscheinung von Medikamenten hohe Praxisrelevanz zusprechen. Weiterhin stellt das während der Fortbildung erworbene Wissen eine Orientierungshilfe im Umgang mit der nicht nur medizinisch, sondern auch juristisch und emotional schwierigen Situation der Aufklärung von älteren Patienten über Einschränkungen der Fahreignung dar. Somit beinhaltet eine Mobilitätsberatung durch den Hausarzt wie keine andere existierende Maßnahme die Möglichkeit, die Fahreignung eines Patienten langfristig zu beobachten und individuell abgestimmte Maßnahmen vorzuschlagen und zu begleiten. Diese kontinuierliche Dokumentation bestehender Leistungseinschränkungen einschließlich vorgeschlagener Kompensationsmechanismen und Anpassungen des Fahrverhaltens garantiert eine möglichst lange Aufrechterhaltung der Mobilität und trägt somit zum Erhalt von Lebensqualität im Alter bei.

Wie die Ergebnisse der Befragung zeigen, sind eigene Fahreignung und die aktive Teilnahme am Straßenverkehr für die Senioren von großer Bedeutung für ihre Lebensqualität. Zugleich zeigten sich jedoch auch mögliche Einschränkungen, durch den Gesundheitszustand der Befragten oder die gleichzeitige Einnahme mehrerer Medikamente. Unabhängig von diesen potenziellen Risikofaktoren schätzen sich die meisten Senioren selbst als weiterhin unbeeinträchtigte, sichere Kraftfahrer ein. Des Weiteren wird Automobilität oft als Garant hoher Lebensqualität und Unabhängigkeit genannt, auf die ungern verzichtet wird. Der Stellung des Hausarztes kommt in dieser

Situation eine besondere Rolle zu, da dieser überwiegend als kompetente Vertrauensperson angesehen wird. Ziel von Mobilitätsberatungen und Aufklärungsaktionen älterer Kraftfahrer sollte daher in erster Linie die Förderung der Wahrnehmung und selbstkritischen Bewertung vorhandener Leistungseinbußen sein. Durch die im Alter zunehmende Multimorbidität und Multimedikation ist die Fahreignung häufig eingeschränkt. Darauf gilt es mit entsprechenden Veränderungen des Mobilitätsverhaltens kompensatorisch zu reagieren. Die Kompetenz und Bereitschaft des Einzelnen, relevante Einschränkungen rechtzeitig zu erkennen und adäquate Kompensationsmaßnahmen zu ergreifen, – ggf. auch eine nicht mehr vorhandene Fahreignung zu akzeptieren – trägt zur individuellen und allgemeinen Verkehrssicherheit entscheidend bei. Das Aufzeigen und die schrittweise Hinführung zur angemessenen Nutzung alternativer Mobilitätsmöglichkeiten (z. B. öffentliche Verkehrsmittel, Fahrrad, etc.) zur Erhaltung der durch eine eigene Mobilitätsfähigkeit gesicherten Lebensqualität sollten ebenfalls Teil einer jeden Beratung sein. Es soll nicht belehrt, sondern zu einer aktiven Auseinandersetzung mit dem Thema Verkehrssicherheit im Alter angeregt werden, mit dem Ziel, die Mobilität so lange wie möglich aufrecht zu erhalten (Kocherscheid et al., 2007).

In diesem Kontext müsste der bereits bestehenden Aufklärungspflicht der behandelnden Ärzte über Risiken einer Erkrankung und damit auch über Tauglichkeitsmängel und ihre Auswirkungen auf die Verkehrssicherheit künftig ausdrücklicher nachgekommen werden. Dieser Beratungsaspekt würde stärker in den Fokus der behandelnden Mediziner rücken, wenn Krankenkassen unterstützend entsprechende spezifische Abrechnungsposten einführen würden. So zeigen die Ergebnisse der vorliegenden Untersuchung, dass Zeitmangel und Probleme bei der Abrechnung aufwändiger Beratungsmaßnahmen die im Einzelfall notwendige Beratung verhindern, da die Entlohnung des Arbeitsaufwandes einen wichtigen Faktor für die generelle Akzeptanz der konzipierten Fortbildung darstellt. Eine flächendeckende Implementierung der Fortbildung darf somit nicht nur die absolute Bereitschaft der Ärzte zur Mitarbeit fordern, sondern sollte zusätzlich mit der Entwicklung eines geeigneten Abrechnungssystems auf höherer Ebene einhergehen, damit eine Vergütung des geleisteten Aufwandes gewährleistet werden kann. Hier wären weiterführende Maßnahmen zu planen, die die Nutzung des Präventionspotenzials einer Mobilitätsberatung optimieren können.

Die Aufnahme einer verkehrssicherheitsrelevanten Beratung als ärztliche Leistung im hausärztlichen Kapitel des neuen EBM¹ könnte einen wichtigen Beitrag für die verkehrssicherheitsrelevante Betreuung von Senioren leisten. Durch die finanzielle Honorierung der ärztlichen Beratung würden sich die Chancen für eine dauerhafte Integration eines mobilitätsbezogenen Bera-

1 Einheitlicher Bewertungsmaßstab

tungsgespräch in den medizinischen Behandlungsalltag erhöhen. Neben diesem finanziellen Anreiz zur Aufklärung könnte für die behandelnden Ärzte im Rahmen der Verordnung ein spezifischer Katalog von verkehrsrelevanten Veränderungen aufgrund von spezifischen Medikamenteneinnahmen und Erkrankungen hilfreich sein. Auch die Information über Medikamentenkombinationen im Zusammenhang mit häufig auftretenden Multimorbiditätsmustern bei Älteren, die hinsichtlich des Mobilitätsaspekts empfehlenswert sind, könnte unterstützend erlebt werden. In diesem Zusammenhang wurde auch der Bedarf an einem Screening-Verfahren geäußert, das ökonomisch eine Erfassung der wichtigsten verkehrssicherheitsrelevanten Funktionsbereiche erlaubt und für den Patienten eine hohe Augenscheinvalidität besitzt.

Modul 1

Medizinisch psychologische Grundlagen des sicheren Verkehrsverhaltens älterer Menschen

In diesem Modul werden medizinische und psychologische Rahmenbedingungen einer sicheren Verkehrsteilnahme älterer Menschen besprochen. U. a. werden Fragen der Wahrnehmungs- und Reaktionsfähigkeit, der Belastbarkeit und Beanspruchung, der intellektuellen Beeinträchtigung sowie der Beeinträchtigung durch Krankheiten beziehungsweise Medikationen thematisiert (speziell Multimorbidität und Multimedikation). Darüber hinaus werden ökonomische Verfahren zur Fahreignungsdiagnostik vorgestellt, Möglichkeiten und Grenzen von Fahrerassistenzsystemen aufgezeigt und Kompensationsmöglichkeiten altersbedingter Einschränkung der Fahreignung thematisiert.

Modul 2

Rechtliche Rahmenbedingungen – Altern und Fahreignung

Die interaktive Auseinandersetzung mit rechtlichen Rahmenbedingungen, die für die Beurteilung der Fahreignung älterer Kraftfahrer gelten, steht im vorliegenden Modul im Vordergrund. Dargelegt werden Fälle, anhand derer folgende Aspekte thematisiert und diskutiert werden: Aufklärungspflicht des Arztes bei Fahruntauglichkeit des Patienten, Unterrichtung Dritter und ärztliche Schweigepflicht sowie Überprüfung der Fahreignung bei Krankheit und der Einnahme von Arzneimitteln. Die Kenntnis gesetzlicher Vorgaben und Spielräume erhöht die Handlungsmöglichkeiten des Arztes.

Modul 3

Verkehrssicherheit und Mobilität älterer Menschen als Themenkomplex im Rahmen der Arzt-Patient-Interaktion

Die Thematisierung der Fahreignung (und gegebenenfalls Mobilitätsbeschränkung) älterer Patienten stellt insbesondere ein kommunikatives Problem dar. Will ein Arzt im Sinne des Patientenwohls deutlich machen, dass er eine weitere Teilnahme am automobilen Straßenverkehr für nicht angezeigt erachtet, stößt er häufig auf Widerspruch und mangelnde Akzeptanz. Durch die Erhöhung der ärztlichen kommunikativen Kompetenz soll die Vermittlung entsprechender Diagnosen und Empfehlungen für den Arzt erleichtert und für den Patienten eher annehmbar gemacht werden. In diesem Modul werden Gesprächs- und Argumentationstechniken (z. B. aktives Zuhören, Vermittlung von Ich-Botschaften, Kosten-Nutzen-Argumentationen etc.) im Rahmen von praktischen Übungen gezielt eingeübt und trainiert. Hierzu werden Rollenspiele eingesetzt ebenso wie die Diskussion von Fallbeispielen und die Entwicklung von Best-Practice-Vorgehensweisen. Darüber hinaus sollen mögliche Schwierigkeiten bei einer Mobilitätsberatung thematisiert und alternative Herangehensweisen erarbeitet werden.

Um eine effiziente überregionale Implementation und Verstetigung der Fortbildungsmaßnahme gewährleisten zu können sowie die generelle Wirksamkeit und den Nutzen der Maßnahme im Hinblick auf die Verbesserung der Verkehrssicherheit älterer Menschen zu überprüfen, wurde die Fortbildung überregional zertifiziert² und implementiert. Parallel dazu wurde ein Evaluationsverfahren entwickelt, um die Einstellung und das Verhalten der älteren Verkehrsteilnehmer sowie die Auswirkungen der Fortbildung auf den ärztlichen Behandlungsalltag zu erfassen.

3. Implikationen für die Praxis

Um den Transfer des in der Fortbildung erworbenen Wissens im Praxisalltag zu gewährleisten und um eine Orientierungshilfe für die Beurteilung der Fahreignung zu schaffen, wäre gemäß den Ergebnissen der vorliegenden Untersuchung die Entwicklung eines Mobilitäts-Screenings für Ärzte sinnvoll, das den Gesundheitszustand, aber auch die fahrpraktischen und kognitiven Kompetenzen der älteren Patienten zur Verkehrsteilnahme erfasst.

Da die Anpassung des Fahrverhaltens eine Einsicht in bestehende Defizite und die Bereitschaft zur Fahrverhaltensänderung voraussetzt, wäre die Ent-

2 Kategorie C/8 Punkte.

wicklung eines Screenings wünschenswert, mit dem älteren Verkehrsteilnehmern augenscheinliche fahrtüchtigkeitsrelevante Einschränkungen demonstriert werden können. Die Güte eines solchen Instrumentes sollte, wie schon angedeutet, an der Augenscheinvalidität für den älteren Verkehrsteilnehmer und nicht zwangsweise an der Kriteriumsvalidität des tatsächlichen Fahrverhaltens gemessen werden, da das Ziel in der Sensibilisierung für persönliche Einschränkungen besteht, nicht in einer normativen Fahrtauglichkeitsbegutachtung. Desweiteren sollte das Screening – laut Aussage der Ärzte – in den Praxisbetrieb integrierbar und somit ökonomisch einsetzbar sein.

Im Rahmen des Screenings sollte eine qualifizierte ärztliche Mobilitätsberatung (eventuell auch als gezielte Fortbildung und Nachschulung) stattfinden, um Möglichkeiten der selbstkritischen Erkennung von Defiziten und deren Kompensation zu bearbeiten und gleichzeitig als Hilfsmittel zu dienen, die Akzeptanz des Verzichts auf das Autofahren bei den betroffenen älteren Patienten zu erhöhen. Die Anwendungsziele eines solchen Screenings könnten wie folgt lauten:

- Übersichtliche und nachvollziehbare Präsentation vorhandener verkehrsrelevanter alters- und krankheitsbedingter Einschränkungen;
- Erzeugung einer realistischen Einschätzung der körperlichen und geistigen Leistungsfähigkeit;
- Anpassung von Leistungsfähigkeit und Fahrverhalten.

Wesentlich ist hier, dass die Vermittlung der Beratungsinhalte stets anschaulich und praxisnah erfolgen: Eine Information über eine Einschränkung der geistigen Leistungsfähigkeit ist nicht besonders anschaulich, wenn den Betroffenen nicht klar wird, was das für die Fahrpraxis konkret bedeutet und welche Maßnahmen in diesem Fall kompensatorisch sinnvoll sind. Im Sinne eines »Train the Trainer« sollten solche Zusammenhänge zwischen altersbedingten Defiziten und dem Fahrverhalten künftig in die Ausbildung angehender Ärzte integriert werden und durch Fortbildungsangebote aufgefrischt und vertieft werden. Solche Beratungsansätze zur Aufrechterhaltung und Förderung der Mobilität stehen gesetzlichen Forderungen nach vorgeschriebenen altersabhängigen Pflichtuntersuchungen der Fahreignung gegenüber, die über die Aufrechterhaltung beziehungsweise Beendigung einer aktiven Verkehrsteilnahme entscheiden sollen, da nicht das Alter per se, sondern die individuelle Befindlichkeit als relevantes Kriterium für spezifische Fahreignung zu werten ist.

Vor diesem Hintergrund scheint eine ärztliche Fortbildung zur Mobilitätsberatung älterer Patienten, die neben medizinisch-psychologischen Grundlagen sowohl rechtliche Rahmenbedingungen der Verkehrsteilnahme älterer Menschen als auch kommunikative Aspekte der Arzt-Patient Interaktion berücksichtigt, den aktuellen und insbesondere den zukünftigen Bedürfnissen und Notwendigkeiten zu entsprechen. Die Umsetzung von Mobilitätsbera-

tungen in der ärztlichen Praxis stellt somit einen wichtigen und zukunftsorientierten Beitrag zur Verkehrssicherheit dar.

Abschließend kann festgehalten werden, dass eine Mobilitätsberatung durch den Hausarzt wie keine andere existierende Maßnahme die Möglichkeit beinhaltet, die Fahrtüchtigkeit eines Patienten langfristig zu beobachten und individuell abgestimmte Maßnahmen vorzuschlagen und zu begleiten. Diese kontinuierliche Dokumentation bestehender Leistungseinschränkungen einschließlich vorgeschlagener Kompensationsmechanismen und Anpassungen des Fahrverhaltens garantiert eine möglichst lange Aufrechterhaltung der Mobilität und trägt somit zum Erhalt von Lebensqualität im Alter bei.

Literatur

- BECKER, S., Berger, R. & Dumbs, M. (2001). Projektgruppe »Perspektiven der Verkehrssicherheitsarbeit mit Senioren«. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Berichte der Bundesanstalt für Straßenwesen, Mensch und Sicherheit* (Heft M 131). Bremerhaven: Wirtschaftsverlag NW.
- BREDER, C. (2001). Aktiv Zuhören und effektiv informieren. Ein Kommunikationstraining für Hausärzte soll die Compliance steigern. *Deutsches Ärzteblatt*, 98 (26), A-1728/B-1482/C-1374.
- COOPER, B. & Sosna, U. (1983). Psychische Erkrankung in der Altenbevölkerung: Eine epidemiologische Feldstudie in Mannheim. *Nervenarzt*, 54, 239-249.
- ENGELN, A. & Schlag, B. (2007). Kompensationsmöglichkeiten und Bewältigungsstrategien im Alter. In B. Schlag (Hrsg.), *Leistungsfähigkeit und Mobilität im Alter* (Schriftenreihe Mobilität und Alter der Eugen-Otto-Butz-Stiftung. Bd. 3, S. 255-273). Köln: TÜV Media GmbH.
- EWERT, U. (2008). Alterskorrelierte Erkrankungen, die die Verkehrsteilnahme beeinträchtigen können. In B. Schlag (Hrsg.), *Leistungsfähigkeit und Mobilität im Alter* (Schriftenreihe Mobilität und Alter der Eugen-Otto-Butz-Stiftung. Bd. 3, S. 181-197). Köln: TÜV Verlag.
- FASTENMEIER, W., Gstalter, H. & Eggerdinger, Ch. (2005). Zweifel an der Fahrtauglichkeit eines Patienten? Schicken Sie ihre Senioren doch mal zum »TÜV«. *Münchener medizinische Wochenschrift, Fortschritte der Medizin*, 147 (40), 40-43.
- FELLHASE, A. (2005). *Mobil im Alter*. Deutscher Juristinnenbund (djb). Zugriff am 18.01.2009 von http://www.djb.de/publikationen/aktuelle-informationen/ai_2005-3/ai_2005-3f2/
- FOZARD, J. L. (2000). Sensory and cognitive changes with age. In K. W. Schaie & M. Pietrucha (Eds.), *Mobility and transportation in the elderly* (pp. 1-61). New York: Springer.
- GAFFRON, P. (2002). *The implementation of walking and cycling policies*. Paper presented at the meeting of STELLA Focus Group 3, Bonn.
- HENNING, J. (2007). Verkehrssicherheitsberatung älterer Verkehrsteilnehmer. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Berichte der Bundesanstalt für*

- Straßenwesen, Mensch und Sicherheit* (Heft M 189). Bremerhaven: Wirtschaftsverlag NW.
- HOFFMANN-Born, H. & Peitz, J. (2006). Ärztliche Aufklärung bei problematischer Fahreignung. *Ärzteblatt Sachsen*, 8, 402-405.
- INNOVATIONS REPORT (2001). *Schwierige ärztliche Gespräche in Klinik und Praxis*. Forum für Wissenschaft Industrie und Wirtschaft. IDEA TV Ges. für kommunikative Unternehmensbetreuung mb.
- KALWITZKI, K. P. (1996). Soziale Kompetenz für Mobilitätsberater. *Verkehrszeichen*, 3 (96), 23-25.
- KOCHERSCHIED, K. & Rudinger, G. (2005). Ressourcen älterer Verkehrsteilnehmer und Verkehrsteilnehmerinnen. In W. Echterhoff (Hrsg.), *Strategien zur Sicherung der Mobilität älterer Menschen* (Schriftenreihe Mobilität und Alter der Eugen-Otto-Butz-Stiftung. Bd. 1, S. 19-42). Köln: TÜV Verlag.
- KOCHERSCHIED, K. (2008). *Konzeption und Evaluation einer ärztlichen Fortbildung zur Mobilitätsberatung älterer Kraftfahrer*. Südwestdeutscher Verlag für Hochschulschriften: Saarbrücken.
- KOCHERSCHIED, K., Rietz, C., Poppelreuter, S., Riest, N., Müller, A., Rudinger, G. & Engin, T. (2007). Verkehrssicherheitsbotschaften für Senioren: Nutzung der Kommunikationspotenziale im allgemeinmedizinischen Behandlungsalldag. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Berichte der Bundesanstalt für Straßenwesen, Mensch und Sicherheit* (Heft M 184). Bremerhaven: Wirtschaftsverlag NW.
- LIMBOURG, M. & Reiter, K. (2001). Das Verkehrsunfallgeschehen im höheren Lebensalter. In A. Flade, M. Limbourg & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (S. 211-227). Opladen: Leske und Budrich.
- MÜLLER, B. H. (2003). *Institutsbericht der Bergischen Universität Wuppertal (ASER)*. Bereich Mensch und Verkehr.
- PEITZ, J. & Hoffmann-Born, H. (2008). *Arzthaftung bei problematischer Fahreignung*. Bonn: Kirschbaum Verlag.
- RUDINGER, G. & Kocherscheid, K. (2007). Verkehrssicherheitsperspektiven für eine in die Jahre kommende Gesellschaft. ADAC/BASt-Symposium am 13. Oktober 2006 in Baden-Baden. In Bundesanstalt für Straßenwesen (BASt) (Hrsg.), *Berichte der Bundesanstalt für Straßenwesen BASt, Mensch und Sicherheit* (Heft M 178). Bremerhaven: BASt.
- STARFIELD, B. (1998). *Primary care. balancing health needs, services and technology*. Oxford: University Press.
- WAHL, H. W., Tesch-Römer, C. & Rott, C. (2000). Vision and cognitive functioning in old age. In B. Silverstone, M. A. Lang & B. Rosenthal et al. (Eds.), *The Lighthouse handbook of vision impairment and vision rehabilitation* (pp. 431-439). New York: Oxford University Press.
- WIPPENBECK, P. (2002). *Einbeziehung älterer Menschen in der hausärztlichen Versorgung. Möglichkeiten und Grenzen*. Diplomarbeit, Hochschule Magdeburg Stendal (FH).

Georg Rudinger & Anton Reiter¹

Zur Evaluation der theoretischen Fahrerlaubnisprüfung²

Zusammenfassung

Vor dem Hintergrund der hochkomplexen Organisation des Fahrerlaubnis- und Prüfwesens in Deutschland erscheint die Einführung einer grundlegend neukonzipierten theoretischen Fahrerlaubnisprüfung mittelfristig nicht möglich. Anknüpfend an die Arbeiten der »Studienstelle für Fahrerlaubnisfragen« des TÜV Rheinland wird ein Verfahren zur testtheoretischen Beurteilung und Optimierung des aktuellen Instruments skizziert. Erste Ergebnisse werden anhand einer Stichprobe von 4.144 Prüfungen für den Führerschein Klasse B berichtet. Zur Einleitung eines kontinuierlichen Verbesserungsprozesses werden folgende, unmittelbar umsetzbare Schritte vorgeschlagen: (1) Einführung von PC-gestützten Theorieprüfungen, (2) Absicherung der testtheoretischen Güte, (3) Entwicklung und Implementierung neuer Prüfungsfragen.

1. Einleitung

Mit dem Forschungsvorhaben »Optimierung der Fahrerlaubnisprüfung« aus dem Jahr 1997 hat die Bundesanstalt für Straßenwesen (BASt) einen wichtigen Anstoß zu konzeptionellen Neuansätzen der theoretischen und praktischen Fahrerlaubnisprüfung gegeben (BASt, 1997). Die kürzlich vorgelegten Analyseergebnisse beziehen sich auf das Gesamtsystem von Fahrausbildung und -prüfung. Sie basieren auf aktuellen Erkenntnissen der Unfall- und Sicherheitsforschung zu den Anforderungen, denen der zumeist jugendliche Fahranfänger im heutigen Straßenverkehr gerecht werden muss, und zeichnen Wege auf, wie unter Verwendung computergestützter Instrumente die methodischen und verfahrenstechnischen Voraussetzungen zur kontinuierlichen Verbesserung der Fahrausbildung und -prüfung geschaffen werden können (Bönninger & Sturzbecher, 2004). Die Umsetzung der Vorschläge dürfte angesichts der hochkomplexen Organisation des Fahrerlaubnis- und Prüfwesens in Deutschland jedoch eine längere Zeitspanne erfordern – zumindest länger als die von den Autoren des BASt-Forschungsberichts angenommenen zwei bis drei Jahre.

1 Erschienen als Rudinger, G., Reiter, A.(2004). Zur Evaluation der theoretischen Fahrerlaubnisprüfung. *Zeitschrift für Verkehrssicherheit* 50 (3), 141-144

2 Wir danken Norbert Hilger für seine Unterstützung.

Der TÜV-Verlag, der seit vielen Jahren mit der Erstellung aller Prüfmedien für die theoretische Fahrerlaubnisprüfung betraut ist, – das sind Prüfbögen auf Papier in 12 Sprachen, ein Audiosystem auf MiniDisc in 7 Sprachen und eine PC-gestützte Prüfung auf der Grundlage der Papierbögen, ebenfalls in 12 Sprachen – war neben vielen anderen Stellen an den Arbeiten im BAST-Forschungsprojekt beteiligt. Angesichts der aufgezeigten Optimierungsmöglichkeiten der theoretischen Fahrerlaubnisprüfung gibt es durchaus Ansatzpunkte für kurzfristig realisierbare Beiträge zur Verbesserung des gegenwärtigen Prüfsystems. Zu diesem Zweck hat der TÜV-Verlag gemeinsam mit dem Zentrum für Evaluation und Methoden der Universität Bonn (ZEM) ein Verfahren entwickelt, das auf der Ebene testinterner Gütekriterien eine empirisch gestützte Evaluation und Optimierung der Prüfungsaufgaben, wie sie noch auf absehbare Zeit zum Einsatz kommen werden, sicherstellen kann. Das methodische Verfahren wurde anhand einer Stichprobe von 4.144 Erstprüfungen für den Führerschein Klasse B aus dem Zeitraum 2003/2004 entwickelt und steht zur Überprüfung der ab dem 1. Juli 2004 eingesetzten neuen Prüffragen und Prüfbögen erstmals zur Verfügung. Es knüpft an frühere Analysen von Prüfergebnissen an, die vor allem vom TÜV Rheinland bereits in den 70er Jahren in seiner »Studienstelle für Fahrerlaubnisfragen« (Schneider, 1977) sowie später vom VdTÜV und der DEKRA (Petzholtz & Erdenkäufer, 2002) vorgenommen wurden. Das vorliegende Verfahren nutzt das vom TÜV-Verlag entwickelte PC-gestützte Prüfungssystem, in das – nachträglich – die Prüfergebnisse der untersuchten 4.144 Prüfungen eingegeben wurden; damit standen anonymisierte Daten von 124.320 Beantwortungen von Einzelfragen aus dem Bereich der theoretischen Fahrerlaubnisprüfung zur Verfügung. Auf dieser Grundlage können die verfügbaren Prüfungsdaten (Ergebnisse und soziodemografische Angaben) stärker als bisher für Evaluationszwecke ausgeschöpft werden.

Nachfolgend wird das methodische Vorgehen einschließlich der Auswertungsverfahren beschrieben und an einigen empirischen Beispielen dokumentiert. Anschließend wird der Wert des hier gewählten Evaluationsansatzes für die kontinuierliche Optimierung des gegenwärtigen Prüfsystems aufgezeigt. Weitere Entwicklungsmöglichkeiten im Sinne des o.g. Forschungsberichts werden zur Diskussion gestellt.

2. Testgütekriterien, Testevaluation und Möglichkeiten der Testoptimierung

Wie jeder andere Leistungstest auch, muss die theoretische Fahrerlaubnisprüfung den allgemein üblichen Testgütekriterien der Objektivität, Reliabilität und Validität genügen. Die Auswertungsobjektivität der theoretischen Fahrerlaubnisprüfung wird durch die Verwendung standardisierter Auswertungsschablonen sichergestellt. Zur Überprüfung der Reliabilität und insbe-

sondere der Validität der theoretischen Fahrerlaubnisprüfung liegen bislang keine belastbaren empirischen Untersuchungen vor. Da die theoretische Fahrerlaubnisprüfung in mehreren Formen und Versionen eingesetzt wird, kommt zudem den Fragen nach der Parallelität der Testbögen, der Äquivalenz von Papier- und Computer-Versionen sowie der Testfairness besondere Bedeutung zu. Die psychometrische Beurteilung der theoretischen Fahrerlaubnisprüfung muss ferner die Besonderheiten eines lehrzielorientierten Tests berücksichtigen. Wie bereits Kroj (Kroj, 1977) zusammenfassend dargestellt hat, wird mit der Fahrerlaubnisprüfung keine differentialdiagnostische Fragestellung verfolgt, sondern es geht vielmehr um den Nachweis, ob das zum Führen eines Kraftfahrzeugs erforderliche Wissen überhaupt vorhanden ist oder eben nicht. Bei der Fahrerlaubnisprüfung erfolgen die Konstruktion der Fragen und die Zusammenstellung der Bögen aus inhaltlichen Erwägungen, eine empirische Selektion besonders trennscharfer Fragen – wie von Bönninger und Sturzbecher (Bönninger & Sturzbecher, 2004) gefordert – wird nicht durchgeführt. Besonderheiten hinsichtlich der Testgütebeurteilung ergeben sich bei lehrzielorientierten Tests vielfach aus einer geringen Varianz der Testwerte. Denn je geringer die Varianz, desto niedriger fallen die klassischen Testkennwerte aus und erweisen sich als ungeeignete Maßstäbe zur Beurteilung von Testverfahren. In solchen Fällen müssen alternative Maße herangezogen werden, wie z.B. Frickes Ü-Koeffizient (Klauer, 1987). Hierbei werden nicht die Zusammenhänge zwischen den Testdaten selbst, sondern die Übereinstimmungen der Urteile verschiedener Testanwender bestimmt.

Während eine inhaltliche Validierung über Expertenratings zufriedenstellend gewährleistet werden kann, fehlt es hinsichtlich der Bewertung der kriteriumsbezogenen Validität an geeigneten Analysen. Die direkte Untersuchung der prognostischen Validität beispielsweise über die Legalbewährung von Fahranfängern erscheint wenig aussichtsreich, da die Fahrerlaubnisprüfung zwischen erfolgreichen Prüflingen, die ja maximal vier Fragen falsch beantwortet haben, kaum mehr zu differenzieren erlaubt. Vielversprechender erscheinen hier experimentelle Ansätze mit gesonderten Stichproben. Um mehr Varianz in die Daten zu bringen, könnte man beispielsweise einer Stichprobe erfolgreicher Prüflinge den kompletten Fragenkatalog zur Bearbeitung vorlegen und anschließend Zusammenhänge mit Kriterien der Fahrbewährung untersuchen. Ein ähnliches Vorgehen ist auch mit bereits verkehrserfahrenen Fahrerlaubnisinhabern denkbar. Mit solchen Untersuchungen ließen sich sowohl Trennschärfen als auch kriteriumsbezogene Validitäten bestimmen.

Über die bloße Erfassung der aktuellen Testeigenschaften hinaus, ist es zudem Aufgabe der Evaluation, die Optimierung der Testqualität und die kontinuierliche Fortentwicklung der Fahrerlaubnisprüfung zu ermöglichen. Hiermit ist zum einen die kurzfristig erreichbare Verbesserung des aktuellen Testmaterials gemeint: Diese Verbesserung beinhaltet im wesentlichen die

sprachlich-inhaltliche Revision von itemanalytisch auffälligen Fragen, die Verwendung geeigneterer Bildmaterials sowie die Sicherung der Parallelität der verschiedenen Bögen. Die Parallelität unterschiedlicher, im Rahmen der theoretischen Fahrerlaubnisprüfung eingesetzter, Testbögen kann weitgehend über die Angleichung der Schwierigkeitsverteilungen hergestellt werden (Lienert & Raatz, 1998). Neue Fragen sollten daher zukünftig erst nach empirischer Vorprüfung in das bestehende Testmaterial eingearbeitet werden. Über eine so gestaltete Evaluation lassen sich aber zum anderen auch Möglichkeiten zur Optimierung der theoretischen Fahrerlaubnisprüfung, beispielsweise über konzeptuelle Erweiterungen wie den Einsatz von bewegtem Bildmaterial oder akustischen Elementen, überprüfen. Auf dieser Grundlage ließe sich eine permanente Verbesserung des seit langem bestehenden und bewährten Verfahrens der theoretischen Fahrerlaubnisprüfung erreichen, die einerseits etablierte Verfahrensprozesse und -inhalte berücksichtigt, und die andererseits Innovationen und Anpassungen nach entsprechender Tauglichkeitsprüfung vorsieht.

3. Empirische Analyse

Um einen aktuellen Einblick in die psychometrische Struktur der Fahrerlaubnisprüfung zu bekommen, hat der TÜV-Verlag aus den Jahren 2003/2004 eine Stichprobe von 4.144 deutschsprachigen Erstprüfungen für den Führerschein Klasse B in elf Prüfstationen aus Nordrhein-Westfalen und Rheinland-Pfalz gezogen. Da diese Stichprobe nicht anhand eines definierten Auswahlmechanismus gewonnen werden konnte, kann leider nicht von Repräsentativität des Datenmaterials ausgegangen werden (Schäffer, 1996). Dennoch vermitteln die Daten einen überzeugenden Eindruck von der Schwierigkeit des Prüfungsmaterials und geben vor allem Anhaltspunkte für die weitere Evaluation.

Itemschwierigkeiten

Der testtheoretische Begriff der Schwierigkeit bezeichnet die relative Häufigkeit, mit der eine Frage richtig gelöst wird. Bild 1 zeigt die Verteilung der Schwierigkeiten aller 577 Fragen, die im untersuchten Zeitraum in den 32 Bögen zum Einsatz kamen. Insgesamt lassen sich die Fragen der Führerscheinprüfung aus testtheoretischer Perspektive als einfach bezeichnen, die meisten Fragen haben eine Schwierigkeit nahe Eins, wobei 45 Fragen (8%) von allen Prüflingen richtig beantwortet wurden. Nach der Vorbereitung durch die theoretische Fahrausbildung lassen sich somit die meisten Fragen der Fahrerlaubnisprüfung richtig beantworten. Dies ist gesamtgesellschaftlich wünschenswert, denn bei Vorliegen entsprechend erlernbarer Kenntnisse soll möglichst vielen Bürgern eine Teilnahme am motorisierten Straßenverkehr ermöglicht werden. Im Durchschnitt haben die Fragen eine Schwierigkeit von

0,91, wobei die Hälfte der Fragen jeweils von mehr als 95% der Prüflinge gelöst wird (Median).

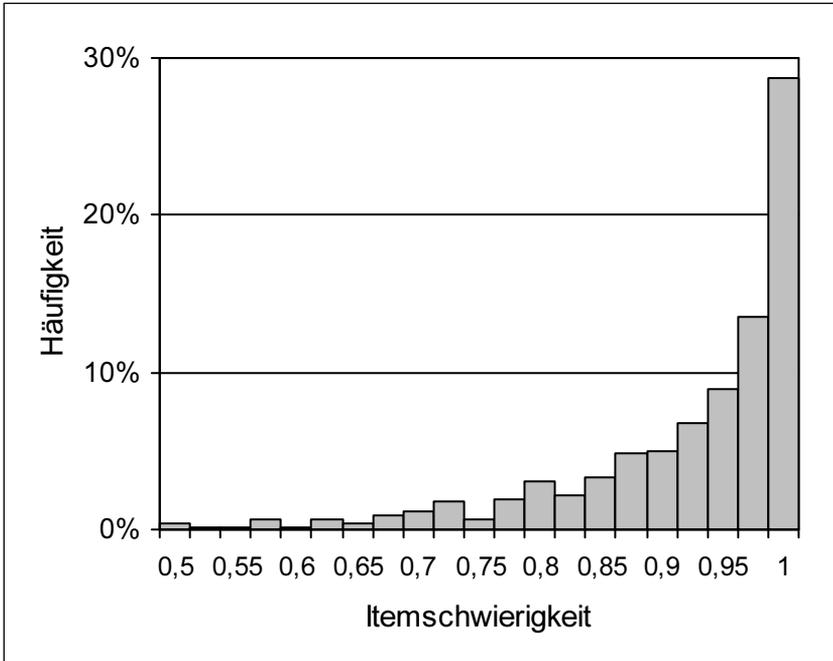


Bild 1: Verteilung aller 577 Itemschwierigkeiten

Neben der Gesamt-Schwierigkeit der Fragen wurden auch sämtliche Antwortalternativen einzeln ausgewertet. Während die Lösungshäufigkeiten der Antwortalternativen bei fast allen Fragen jeweils ähnlich sind, ergaben sich bei den offensichtlich als schwierig einzustufenden Fragen signifikante Auffälligkeiten für einzelne Antwortalternativen. Bild 2 zeigt die Schwierigkeiten der 30 Fragen eines einzelnen Prüfbogens.

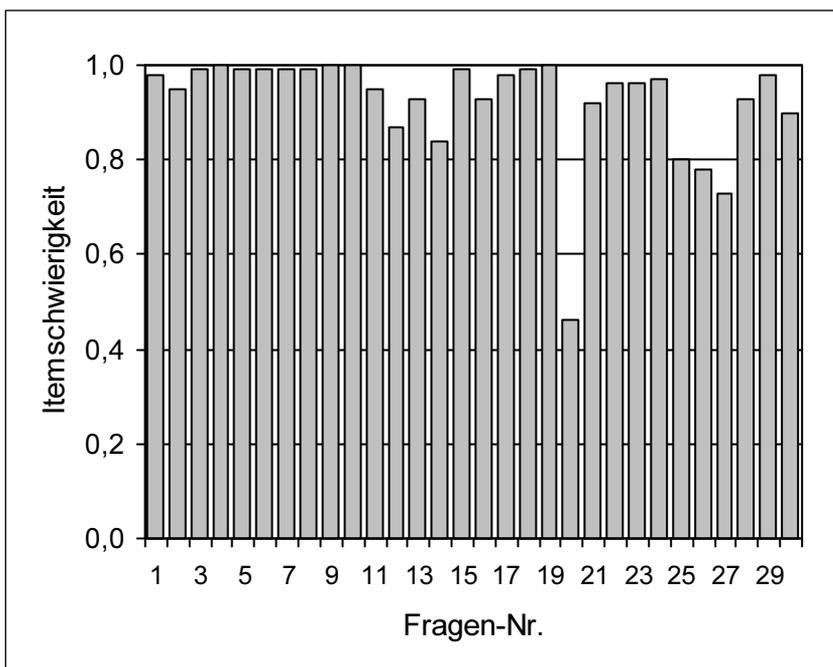


Bild 2: Schwierigkeiten eines Bogens

Auffällig ist, dass die Frage Nr. 20 die mit Abstand schwierigste Frage dieses Prüfungssets ist. Eine genauere Betrachtung dieser Frage (es ist die Frage Nr. 1.5.01-004 des amtlichen Fragenkataloges) zeigt Folgendes (Tabelle 1):

Tabelle 1: Schwierigkeiten der Antwortalternativen einer Frage

Was muss regelmäßig gewartet werden, um zu hohem Kraftstoffverbrauch und übermäßigen Schadstoffausstoß zu vermeiden?	Schwierigkeit
Antwort 1: Motor-Luftfilter	0,90
Antwort 2: Vergaser oder Einspritzanlage	0,96
Antwort 3: Zündanlage	0,48

Für die hohe Quote der Falschbeantwortung ist also ganz wesentlich Antwortalternative 3 verantwortlich, 52% der Prüflinge kreuzen die Antwort 3 fälschlicherweise nicht an. Diese Aussage gibt dem Expertenteam, das die Prüfungsfragen erarbeitet, deutliche Hinweise, welche Antwortalternativen kritisch geprüft und möglicherweise auch ersetzt werden sollten. Dabei ist zudem zu berücksichtigen, dass bei solch auffälligen Falschbeantwortungen möglicherweise Defizite bei der Prüfungsvorbereitung bestehen, z. B. weil

die Thematik »Kraftstoffverbrauch« vielleicht nicht so wichtig genommen wird.

Profile gemäß Bild 2 liegen für alle derzeit eingesetzten 32 Testbögen vor. In vielen Fällen zeigt sich, dass Fragen aus dem klassenspezifischen Zusatzstoff (hier Klasse B) – also Fragen der Nr. 24-30 – besonders häufig falsch beantwortet werden. Auch dies mag mit Defiziten in der Prüfungsvorbereitung und/oder mit der Einstellung der Probanden, dass die 23 Fragen des Grundstoffs »wichtiger« sind, zusammenhängen. Die durchschnittliche Schwierigkeit des Grundstoffs liegt bei 0,94, die des Zusatzstoffs bei 0,85. Ein typisches Profil ist das des in Bild 3 wiedergegebenen Bogens, bei dem die Fragen 25, 26, 28 und 30 als besonders schwierig auffallen.

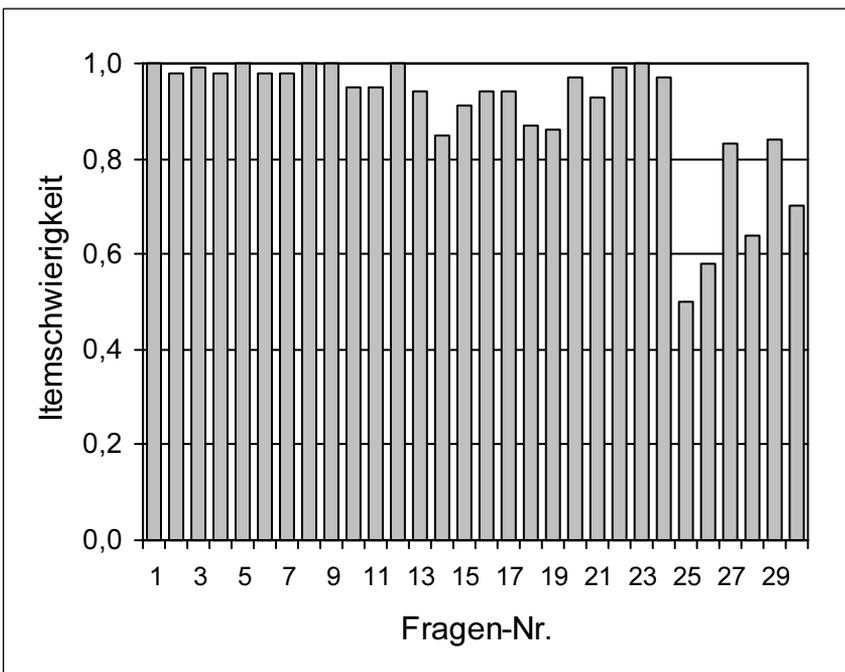


Bild 3: Schwierigkeiten eines Bogens

Die Ergebnisse dieser Analyse ermöglichen auch die Erstellung einer Rangliste der in allen Prüfungen verwendeten 577 Fragen nach dem Grad der Falschbeantwortung. Danach lassen sich die 10 schwierigsten Fragen den folgenden Stichworten zuordnen, von denen 8 aus dem Bereich »Zusatzstoff« kommen. Es sind Fragen zu

- Anhängelast,
- Sonntagsfahrverbot,
- Betriebserlaubnis,
- Fahren bei Gefälle (Kupplung),

- Lichthupe,
- Vorteile von ABV und ASR,
- Verkehrszeichen Steigung/Gefälle,
- Umweltschutz (regelmäßige Wartung der Zündanlage).

Die beiden letzten Stichworte beziehen sich auf Fragen aus dem »Grundstoff«. Unter allen Verkehrszeichen führt offensichtlich das Verkehrszeichen »Steigung/Gefälle« zu besonders häufiger Falschbeantwortung. Hier stellt sich auch die Frage, in wie weit die sprachliche Differenzierung der jugendlichen Führerscheinbewerber ausreicht, um die Begriffe Steigung und Gefälle in Verbindung mit dem dargestellten Zeichen eindeutig zuzuordnen (PISA?).

Diese und ähnliche vielfach für das Bestehen oder Nichtbestehen verantwortlichen Fragen stammen aus Bereichen, die nicht mit »Verhalten in kritischen Verkehrssituationen« beschrieben werden können. Insbesondere zeigt sich, dass in der Rangliste der Falschantworten die Vorfahrtsfragen erst ab Rang 130 vorkommen. Existenziell bedeutsame Verkehrsregeln werden somit offensichtlich von (fast) allen Prüflingen gewusst.

Trennschärfe

Der wichtigste Itemkennwert der klassischen Testtheorie ist die Trennschärfe. Die Trennschärfe einer Frage ist die Korrelation zwischen den Punkten dieser Frage und dem jeweiligen Testwert. Je höher die Trennschärfe, desto besser kann mit einer Frage zwischen Könnern und Nichtkönnern unterschieden werden. Würde man die Bewertungskriterien der klassischen Testtheorie heranziehen (und es gibt gute Gründe, dies im vorliegenden Fall nicht zu tun; siehe oben bzw. (Kroj, 1977)), müssten die Trennschärfen der Items der theoretischen Fahrerlaubnisprüfung als sehr niedrig bezeichnet werden (Linert & Raatz, 1998). Der Mittelwert aller Trennschärfen liegt bei 0,18. Diese niedrigen Trennschärfekoeffizienten können aber als direkte Konsequenz der geringen Schwierigkeiten angesehen werden, sowie auch der unbekanntenen Dimensionalität des Verfahrens. Ähnlich verhält es sich mit der Konsistenzschätzung (Cronbachs α), die insgesamt mit 0,63 niedrig ausfällt.

Fehlerverteilung

Tabelle 2 zeigt, wie sich die Fehlerpunkte aller hier berücksichtigten Testbögen insgesamt verteilen. Über 20% der Prüflinge machen keinen Fehler, etwa die Hälfte erhalten bis zu 6 Fehlerpunkte. Ab zehn Fehlerpunkten wird die Prüfung als nicht bestanden gewertet, die Durchfallquote liegt also bei 32%.

Tabelle 2: Verteilung der Fehlerpunkte

	<i>Fehlerpunkte</i>	<i>Prozent</i>	<i>Kumulierte Prozente</i>
bestanden	0	21%	21%
	2	3%	24%
	3	12%	36%
	4	5%	42%
	5	5%	46%
	6	7%	53%
	7	6%	59%
	8	4%	64%
	9	5%	68%
nicht bestanden	10	4%	72%
	11	4%	76%
	12	3%	79%
	13	3%	81%
	14	2%	84%
	15	2%	86%
	16	2%	88%
	17	2%	89%
	18	1%	91%
>18	9%	100%	

Von Bedeutung für die Parallelität der neuen Prüfungsbögen ist insbesondere die Möglichkeit, die Durchfallquote aus den Schwierigkeiten vorherzusagen. Diese Prognose wurde hier mit einem exponentiellen Verteilungsmodell vorgenommen, das sich in einer Kreuzvalidierung mit einem mittleren absoluten Fehler von 1% als hierfür geeignet erwiesen hat. Dies deutet darauf hin, dass auch nach dem Austausch von Fragen die Gerechtigkeit des Verfahrens über die Konstanzhaltung der Schwierigkeiten gesichert werden kann. Hierdurch kann einer ungleichen Behandlung durch verschiedenen schwere Prüfungsbögen, die früher zum Ausschluss entsprechender Bögen führte (Schneider, 1977), vorgebeugt werden.

4. Ausblick

Der BAST-Forschungsbericht schließt u.a. mit der Empfehlung, kurzfristig 10 Prüfungszentren zur Durchführung der theoretischen Fahrerlaubnisprüfung – vorwiegend in Großstädten – einzurichten, in denen PC-Prüfungen angeboten werden. Die auf diese Weise erhaltenen Daten können ohne weiteren Aufwand einer statistischen Analyse zugeführt werden. Das vom TÜV-Verlag und dem ZEM entwickelte methodische Verfahren liefert zeitnah konkrete Ergebnisse zur Qualität der eingesetzten Testfragen, die umgehend wieder für die Entwicklung weiterer Paralleltests sowie auch für die Optimie-

rung einzelner Fragen (Texte und Bilder) im Rahmen der Theorieprüfung genutzt werden können. Auf diese Weise wird ein kontinuierlicher Verbesserungsprozess eingeleitet, wie er Teil jedes Qualitätsmanagementsystems ist. Mit den so gewonnenen größeren Stichproben lassen sich auch detailliertere Einzeluntersuchungen durchführen, so z.B. zur Fairness unterschiedlicher sprachlicher Versionen der Testbögen oder zu soziodemographischen Merkmalen der Testabsolventen. Die notwendigen technisch-methodischen Voraussetzungen für die Umsetzung eines solchen Vorgehens (PC-gestütztes Prüfsystem, Datenbanken, Statistikprozeduren) stehen unmittelbar zur Verfügung und können schon ab dem 01. Juli 2004 zum Einsatz kommen.

Neben der Analyse »echter« Prüfungsfragen können über das entwickelte methodische Vorgehen auch neue Fragen, ja völlig neue Fragetypen/-Visualisierungsmethoden in einer Art Vortest untersucht und evaluiert werden – eine Methode, die in England und Schweden längst zum Standard gehört.

Der unmittelbar umsetzbare »Masterplan« zur Optimierung der theoretischen Fahrerlaubnisprüfung sieht demnach drei Schritte vor:

1. Einführung von PC-gestützten Theorieprüfungen in großer Zahl.
2. Analyse und Eliminierung von Schwachpunkten aktuell eingesetzter Fragen (sprachliche Verbesserung, Angleichung der Testschwierigkeiten, Erhöhung der Testfairness).
3. Entwicklung und abgesicherte Implementierung neuer Prüfungsfragen, speziell unter Einbeziehung des vorliegenden Erfahrungswissens der bislang am Prozess der Fragenentwicklung beteiligten Experten.

Durch die Einführung einer solchen permanenten formativen Evaluation des Prüfverfahrens lassen sich darüber hinaus eine Vielzahl von Maßnahmen auch zur summativen Evaluation der Methodik realisieren.

Eine schnelle Umsetzung des vorgeschlagenen Vorgehens ist somit der Garant für eine umfassende Qualitätssicherung im Bereich der theoretischen Fahrerlaubnisprüfung.

Literatur

- BUNDESANSTALT für Straßenwesen. (1997). *Sicherheitsforschungsprogramm*. Bergisch-Gladbach: BAST.
- BÖNNINGER, J. & Sturzbecher, D. (2004). *Qualität der Fahrerlaubnisprüfung. Ein Reformvorschlag für die theoretische Fahrerlaubnisprüfung. Abschlussbericht zum Projekt »Optimierung der Fahrerlaubnisprüfung« im Auftrag der Bundesanstalt für Straßenwesen*. (Unveröffentlichter, interner Projektbericht).
- LIENERT, G. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- KLAUER, K.J. (1987). *Kriteriumsorientierte Tests*. Göttingen: Hogrefe.

- KROJ, G. (1977). Führerscheinprüfung – Differentialdiagnose oder lehrzielorientierte Messung? *Mensch-Fahrzeug-Umwelt*, 4, 35-47.
- PETZHOLTZ, W. & Erdenkäufer, B. (2002). *Schwierigkeitsanalyse von Aufgaben der Theoretischen Fahrerlaubnisprüfung*. Unveröffentlichte DEKRA-Studie.
- SCHÄFFER, K.-A. (1996). Planung von Stichprobenerhebungen. In E. Erdfelder, R. Mausfeld, T. Meiser & G. Rudinger (Hrsg.), *Handbuch quantitative Methoden* (S. 23-35). Weinheim: Psychologie Verlags Union.
- SCHNEIDER, W. (1977). Bestandsaufnahme und Entwicklungsaspekte der Fahrerlaubnisprüfung. *Mensch-Fahrzeug-Umwelt*, 4, 17-34.

Renate Soellner, Stefan Huber, Norbert Lenartz & Georg Rudinger

Gesundheitskompetenz

Im Projekt »*Gesundheitskompetenz – Modellentwicklung und Validierung*«¹ wird ein Kompetenzstrukturmodell für die Domäne Gesundheit mit dem Ziel erarbeitet, darauf aufbauend Messinstrumente zur Erfassung der Gesundheitskompetenz von Kindern und Jugendlichen zu entwickeln. So dann sollen Kompetenzniveaus modelliert werden, wobei sich das Vorgehen an Wilson (2005) anlehnt. Das Modell gibt Aufschluss über die Fähigkeiten und Fertigkeiten, die das Konstrukt der Gesundheitskompetenz konstituieren und liefert erste Informationen über die Beziehungen der einzelnen Strukturkomponenten zueinander. Das Ziel besteht darin, unterschiedliche qualitative wie auch quantitative Ausprägungen der Gesundheitskompetenz zu erfassen. Zur Messung der Kompetenzkomponenten sollen Szenarien eingesetzt werden, ähnlich der Messung aggressiven Verhaltens bei Kindern (Lösel & Bliessner, 2003). Die so entwickelten Szenarien stellen zusammen mit Aufgaben zur Erfassung von Grundfertigkeiten (*literacy, numeracy*) und des Systemwissens einen zentralen Schritt bei der Entwicklung von bereichs- und stufenspezifischen Gesundheitskompetenztests für den schulischen Kontext dar.

Konzepte der Gesundheitskompetenz

Dem Begriff der Gesundheitskompetenz (*health literacy*) wird seit der inhaltlichen Neudefinition und Erweiterung durch die Weltgesundheitsorganisation vor einem Jahrzehnt (WHO, 1999; Nutbeam, 2000) viel Aufmerksamkeit aus den verschiedensten Bereichen der Wissenschaft und Gesellschaft zuteil (Soellner, Huber, Lenartz & Rudinger, ang). Er findet sich in den Statuten der WHO ebenso wie in Strategiepapieren der Europäischen Union (European Patients Forum (EPF)-Conference, 2008). Maßnahmen zur Förderung der Gesundheitskompetenz werden in der Schule (z.B. GeKo^{Kids}, »Gesundheitskompetenz bei Kindern in der Schule«; Splieth, Müller, Kohlmann, John & Hoffmann, 2007) ebenso diskutiert wie im Arbeitsleben (Kriegesmann, Kottmann, Masurek & Nowak, 2005). Dabei besteht bis heute kein

1 Dieses Projekt wird im DFG-Schwerpunktprogramm 1293 »Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen« gefördert.

wissenschaftlicher Konsens bezüglich der Frage, wie »Gesundheitskompetenz« zu definieren ist und welche Fähigkeiten und Fertigkeiten diese Kompetenz konstituieren (Baker, 2006; EPF-Conference, 2008).

Generell lassen sich zwei verschiedene Arten des Herangehens an den Begriff der Gesundheitskompetenz unterscheiden (Pleasant & Kuruvilla, 2008). Forschungsarbeiten des *klinischen Ansatzes* verstehen Gesundheitskompetenz als eine Art »Gesundheits-Alphabetisierung«, welche insbesondere Grundfertigkeiten wie Lesen, Verarbeiten und Verstehen von gesundheitlichen Informationen im Kontext der Arzt-Patientenbeziehung oder im Kontakt mit dem Gesundheitswesen umfasst (Tones, 2002; Parker, Baker, Williams & Nurss, 1995). Gesundheitskompetenz wird in diesem Rahmen definiert als »*the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions*« (Ratzan & Parker, 2000, S. vi) oder noch spezifischer als »*the ability to read and comprehend prescription bottles, appointment slips, and the other essential health-related materials required to successfully function as a patient*« (American Medical Association, 1999, S. 552).

Der in einem zentralen Artikel von Nutbeam (2000) erstmals ausformulierte *public health-Ansatz* zur Gesundheitskompetenz erweitert diese Definition um einen aktiven und konstruktiven Umgang mit gesundheitlicher Information und versteht Gesundheitskompetenz als »*the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health*« (WHO, 1998, S.10).

Eine im deutschen Sprachraum viel beachtete Definition versteht Gesundheitskompetenz allgemein als die Fähigkeit des Einzelnen, im täglichen Leben Entscheidungen zu treffen, die sich positiv auf die Gesundheit auswirken (Kickbusch, Maag & Saan, 2005). Diese Entscheidungen beziehen sich auf den häuslichen Bereich, die Gesellschaft allgemein, den Arbeitsplatz, das Gesundheitssystem, den Angebotsmarkt und politische Handlungsfelder. Abel und Bruhin (2003) konzipieren Gesundheitskompetenz vorwiegend als wissensbasierte Kompetenz, zu der sowohl alltagspraktisches als auch spezifisches Wissen gehört und die dem Individuum eine gesundheitsförderliche Lebensführung ermöglicht. Dabei wird Gesundheitskompetenz als eine gesellschaftlich vermittelte Kompetenz angesehen (Abel & Walter, 2002), die primär über Kultur, Bildung und Erziehung erlernt wird (Abel & Bruhin, 2003).

Ungeachtet der erhöhten Aufmerksamkeit, die dem Begriff Gesundheitskompetenz von Seiten der Wissenschaft, der Gesellschaft und der Politik zuteil wird, besteht kein Konsens darüber, was genau unter dem Begriff zu verstehen ist. Baker (2006) stellt in diesem Zusammenhang fest: »*The term <health literacy> has come to mean different things to various audiences and has become a source of confusion*« (S. 878). So stellt sich die Frage, inwie-

fern »Gesundheitskompetenz« ausschließlich kontextspezifische kognitive Fähigkeiten und Fertigkeiten oder auch motivationale Aspekte und Persönlichkeitsmerkmale beinhaltet, inwieweit die allgemeine Intelligenz eine Rolle spielt und wie einzelne Komponenten dieser Kompetenz miteinander in Beziehung stehen. In kompetenztheoretischen Begriffen ausgedrückt fehlt ein Kompetenzstrukturmodell der Gesundheitskompetenz, das über die Komponenten des Begriffs, deren Beziehungen untereinander sowie deren Einbettung zu anderen Konstrukten Auskunft gibt. Darüber hinaus erlaubt die Untersuchung der Beziehung des Konstrukts zu gesundheitsbezogenen Outcomes Rückschlüsse auf dessen Sinnhaftigkeit oder Validität. Erst ein systematisch entwickeltes Strukturmodell der Gesundheitskompetenz stellt eine solide Basis für die Modellierung von Kompetenzniveaus und die Entwicklung von geeigneten Messinstrumenten dar.

Modelle der Gesundheitskompetenz

Ausgehend von den unterschiedlichen Definitionen der Gesundheitskompetenz wurden verschiedene Modelle vorgeschlagen, die entweder einzelne Komponenten der Gesundheitskompetenz und ihre Verknüpfungen abbilden (Nutbeam, 2000; Schulz & Nakamoto, 2005; Kriegesmann et al. 2005) oder aber Bereiche beschreiben, in denen gesundheitskompetentes Handeln sichtbar werden kann (Zarcadoolas, Pleasant & Greer, 2006; Kickbusch, 2006). Das am häufigsten zitierte Modell der Gesundheitskompetenz stammt von Nutbeam (2000) und stellt diese als ein komplexes Konstrukt aus Eigenschaften und Fähigkeiten dar, die es Menschen erlaubt, im Gesundheitssystem erfolgreich zu navigieren. In seinem Stufenmodell beschreibt Nutbeam drei aufeinander aufbauende Formen der Gesundheitskompetenz (siehe Abb. 1).

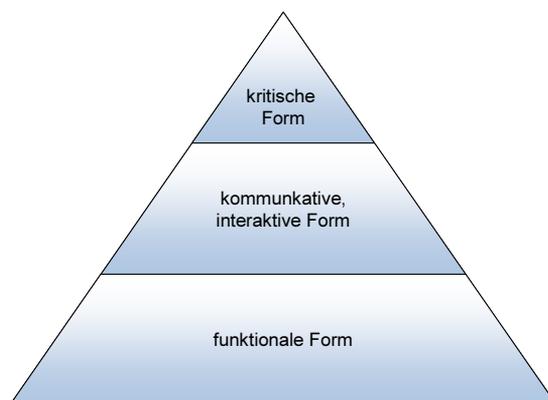


Abbildung 1: Stufenmodell der Gesundheitskompetenz von Nutbeam (2000).

Die *funktionale Form* der Gesundheitskompetenz bildet die unterste Stufe des Modells. Sie beschreibt basale kognitive Fertigkeiten wie Lesen und Schreiben und ermöglicht dem Individuum ein grundlegendes Verständnis gesundheitsrelevanter Informationen. Die zweite Stufe dieses Modells bildet die *kommunikative, interaktive Form* der Gesundheitskompetenz und umfasst fortgeschrittene kognitive und soziale Fertigkeiten, die nötig sind, um eine aktive Rolle im Gesundheitssystem einnehmen zu können. Die oberste Stufe des Modells von Nutbeam ist die sogenannte *kritische Form* der Gesundheitskompetenz. Sie beschreibt die Fähigkeit zur Bewertung von sowie zur Entscheidung für verschiedene gesundheitsrelevante Angebote und Leistungserbringer, wie auch die Fähigkeit zur Compliance mit solchen Angeboten. Es geht dabei um eine kritische Auseinandersetzung mit gesundheitsrelevanten Informationen und dem Gesundheitssystem.

Ein wesentlicher Kritikpunkt an dem Modell ist dessen mangelnde begriffliche Schärfe (Tones, 2002): Die verwendeten Begriffe wie »interaktive« oder »kritische Gesundheitskompetenz« sind nur vage definiert und werden inhaltlich eher in heuristischer Art und Weise umrissen. Zudem wurde der stufenweise Aufbau des Modells mit seinen drei Kompetenzdimensionen bislang empirisch nicht überprüft. Lediglich für die unterste Stufe des Modells, der funktionalen Gesundheitskompetenz, liegen umfassende Forschungsarbeiten vor. Diese sind überwiegend eher dem klinischen Ansatz (s.o.) verpflichtet, dessen Gesundheitskompetenzdefinition inhaltlich weitgehend mit der funktionalen Form aus dem Modell von Nutbeam (2000) übereinstimmt. In neuerer Zeit wurden erste Versuche unternommen, für die interaktive und kritische Form der Gesundheitskompetenz Messinstrumente zu entwickeln (Steckelberg, Hülfenhaus, Kasper, Rost & Mühlhauser, 2007; Ishikawa, Nomura, Sato, & Yano, 2008). Bemühungen, das Gesamtmodell in seiner Struktur abzubilden, sind bisher jedoch nicht unternommen worden.

Schulz und Nakamoto (2005) legen ein Modell zur Gesundheitskompetenz vor, das dieses als multidimensionales Konstrukt konzipiert, bei dem zunehmend komplexere kognitive Fähigkeiten und Fertigkeiten aufeinander aufbauen. Der Grad der Gesundheitskompetenz nimmt zu, indem ein Individuum von basalen Lese- und Schreibfertigkeiten hin zu einem flexiblen und adaptiven Umgang mit neuem Wissen gelangt. Gesundheitskompetenz stellt in diesem Modell die Fähigkeit zum inneren Wissensmanagement und zur Entscheidungsfindung dar. Die Umsetzung in Handeln kann jedoch individuell variieren. Die Autoren betonen in diesem Zusammenhang, dass Gesundheitskompetenz nicht mit Gesundheitsverhalten gleichzusetzen sei. Jemand, der über fundiertes Wissen über die gesundheitsschädigenden Folgen des Rauchens verfüge und trotzdem rauche, sei nicht zwingendermaßen weniger gesundheitskompetent als ein Nichtraucher.

Die Diskrepanz zwischen vorhandenem Wissen, Fertigkeiten und Fähigkeiten und dem tatsächlichen Handeln von Individuen wird im Modell von

Kriegesmann et al. (2005) thematisiert. Dazu betten die Autoren das Konzept der Gesundheitskompetenz in das Bochumer Modell für Handlungskompetenz (Staudt, Kailer, Kriegesmann, Meier, Stephan & Ziegler, 2002) ein. Gesundheitskompetenz setzt sich in dieser Konzeption aus den Komponenten Handlungsfähigkeit und Handlungsbereitschaft zusammen. Zur Handlungsbereitschaft werden subjektive Situationswahrnehmungen und -beurteilungen sowie die individuelle Motivstruktur von Individuen gezählt. Das Modell ist explizit im Kontext beruflicher Beschäftigung verortet und betrachtet Gesundheitskompetenz als Ressource für den Erhalt und die Förderung individueller Beschäftigungsfähigkeit.

Ein weiteres, viel beachtetes Modell zur Gesundheitskompetenz stammt von Kickbusch (2006) (siehe Abb. 2). Aufbauend auf dem Stufenmodell von Nutbeam beschreibt Kickbusch fünf Handlungsfelder, in denen Gesundheitskompetenz bedeutsam ist, und benennt entsprechend Verhaltensweisen, die als hoch oder niedrig kompetent einzustufen sind. Basierend auf den theoretischen Überlegungen Nutbeams werden von Kickbusch auf diese Weise Kontexte beschrieben, in denen die einzelnen Stufen der Gesundheitskompetenz handlungsrelevant sind. Eine empirische Überprüfung des Modells hat jedoch bisher nicht stattgefunden.

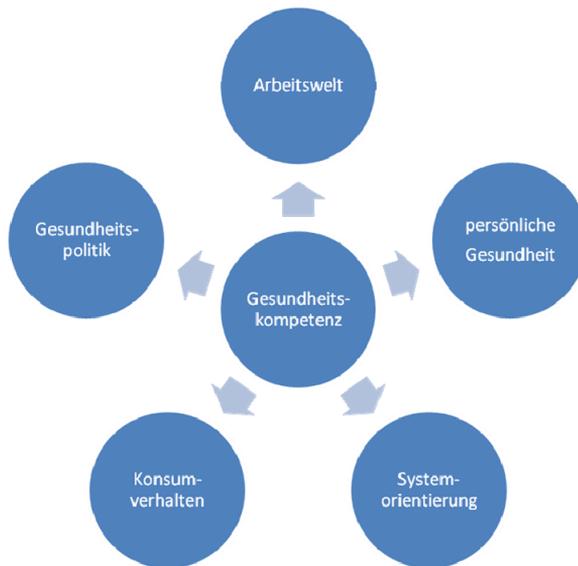


Abbildung 2: Handlungsbereiche der Gesundheitskompetenz nach Kickbusch (2006)

Insgesamt handelt es sich beim Thema Gesundheitskompetenz um ein Forschungsfeld, das von Wissenschaftlerinnen und Wissenschaftlern aus unterschiedlichen Disziplinen bearbeitet wird. Damit gehen unterschiedliche

Schwerpunktsetzungen und Konzeptualisierungen der einzelnen Kompetenzkomponenten einher. So wird Gesundheitskompetenz wahlweise als wissensbasierte Kompetenz oder als Handlungskompetenz verstanden. Entsprechend dem Kompetenzbegriff nach Tippelt, Mandl und Straka (2003) lassen sich Parallelen herstellen zur methodisch-instrumentellen Kompetenz sowie zur sozialen, kommunikativen Kompetenz und auch zum Basiswissen. Die Motivation hingegen wird lediglich in der WHO-Definition sowie in dem Modell von Kriegesmann et al. (2005) angesprochen.

Trotz der unterschiedlichen Schwerpunktsetzungen sowohl bei der Konzeptualisierung des Begriffs der Gesundheitskompetenz als auch bei der Modellbildung des Konstrukts Gesundheitskompetenz lassen sich Kernelemente erkennen, die von der Mehrheit der gängigen Definitionsansätze und Modelle übereinstimmend als zentral erachtet werden (Soellner et al., ang.). Hierzu gehören zum einen grundlegende Fertigkeiten (Lesen, Textverständnis und Rechnen) und einschlägiges Wissen, zum anderen aber auch komplexere Fähigkeiten wie die Wissensorganisation sowie die Fähigkeit zur Bewertung und zur Entscheidungsfindung. Eine umfassende, theoretisch fundierte Modellbildung steht bislang noch aus, und die in den bisher veröffentlichten Modellen postulierte Struktur wurde jeweils nicht systematisch untersucht.

Instrumente zur Erfassung der Gesundheitskompetenz

Die am häufigsten verwendeten Instrumente zur Erfassung von *health literacy* im Sinne des klinischen Paradigmas (Pleasant & Kuruvilla, 2008) sind der *Wide Range Achievement Test* (WRAT), der *Rapid Estimate of Adult Literacy in Medicine* (REALM) und der *Medical Terminology Achievement Reading Test* (MART), die alle die Lesefähigkeit im engeren Sinn erfassen. Der *Test of Functional Health Literacy in Adults* (TOFHLA) ist als Verständnistest konzipiert und prüft an alltagsnahen Beispielen, wie gut die Probanden vorgegebene gesundheitsrelevante Informationen verstehen und nutzen können (Andrus & Roth, 2002; Baker, 2006; Maag, 2007). Ein etwas umfassenderer Test, die *Health Activities Literacy Scale* (HALS; Rudd, Kirsch & Yamamoto, 2004), erfasst Gesundheitsverhalten sowie einige Kompetenzen, wie den Erwerb und die Anwendung von Wissen im Umgang mit Krankheiten oder die Fähigkeit, sich im System zurechtzufinden und seine Rechte als Patient bzw. Bürger zu kennen. Insgesamt sind die bislang entwickelten Instrumente überwiegend auf die Erfassung von Lesefähigkeit im medizinischen Kontext oder aber auf spezifische Wissenskomponenten ausgerichtet (Soellner et al., ang.). Komplexere Fähigkeiten, wie gesundheitsrelevante Entscheidungen zu treffen oder sich im Gesundheitssystem zurechtzufinden, blieben bislang weitestgehend ausgespart. Allerdings wurden in den letzten Jahren erste Versuche unternommen, Messinstrumente zur Erfas-

sung der kommunikativen und kritischen Gesundheitskompetenz zu entwickeln (Steckelberg et al., 2007; Ishikawa et al., 2008).

Der *Critical Health Competence Test* (CHC; Steckelberg et al., 2007) wurde für Studierende eines Studiengangs zur evidenzbasierten Medizin konzipiert. Dieser raschskalierte Fragebogen erfasst kritische Gesundheitskompetenz im Sinne von Nutbeam (2000) mit insgesamt 72 Items in vier Kompetenzbereichen. Diesen Kompetenzbereichen werden jeweils vier medizinische Inhaltsbereiche und vier unterschiedliche Szenarien zugeordnet. Der Fragebogen weist eine sehr gute Reliabilität ($\alpha=.91$) auf, kann aber aufgrund des kognitiven Anspruchsniveaus der Items nur bei Personen eingesetzt werden, die über mindestens 10 Jahre Schulbildung verfügen. Ishikawa et al. (2008) veröffentlichten eine Studie, bei der die beiden Formen der kommunikativen und kritischen Gesundheitskompetenz (3 bzw. 2 Items) aus der Definition der WHO abgeleitet und für japanische Büroangestellte operationalisiert wurden. Ein Instrument zur Erfassung der Gesundheitskompetenz als Handlungskompetenz wurde bislang lediglich am Institut für Sozial- und Präventivmedizin der Universität Zürich entwickelt. Basierend auf den Ergebnissen einer telefonischen Befragung der Schweizer Bevölkerung zum Gesundheitssystem der Zukunft (Future Patient Projekt; Schweizerische Gesellschaft für Gesundheitspolitik, 2003) und dem OECD-Modell für Schlüsselkompetenzen (DeSeCo, 2005) konstruierten Schweizer Forscher einen Fragebogen mit 150 Items, der gesundheitsbezogene Handlungskompetenz in den Bereichen Wissen, Haltung, Werte und Verhalten erfasst (Wang & Schmid, 2007). Bisher existieren noch keine Veröffentlichungen, die Auskunft über die Operationalisierung der postulierten Kompetenzbereiche oder die psychometrischen Eigenschaften des Fragebogens geben. Auch das Instrument selbst ist derzeit noch nicht öffentlich zugänglich.

Obwohl in den letzten Jahren im Bereich der Instrumentenentwicklung zur Erfassung der *health literacy* bzw. von Teilaspekten der Gesundheitskompetenz erste Anstrengungen unternommen wurden, ist das grundlegende Problem, ein Instrument zu entwickeln, das eines der vorgestellten Modelle in seiner Gesamtheit erfasst und die Teilfacetten angemessen berücksichtigt, noch nicht befriedigend gelöst.

Am Stand der Entwicklung der Messinstrumente wird auch der Zustand der Begriffs- bzw. Konstruktbildung des Gesundheitskompetenzkonzeptes deutlich. Solange es an theoretisch verankerter oder empirisch fundierter Forschung zur Konstruktbildung mangelt, lassen sich keine inhaltlich übereinstimmenden Erhebungsinstrumente konstruieren. Um das Konzept der Gesundheitskompetenz umfassend empirisch abbilden zu können, wird ein Kompetenzstrukturmodell benötigt, welches diejenigen Fertigkeiten und Teilkompetenzen ausweist, die notwendig sind, um gesundheitskompetent entscheiden und handeln zu können (vgl. Soellner et al., ang.).

Die Relevanz des Konzepts der Gesundheitskompetenz für den Bereich Bildung zeigt sich u.a. darin, dass der Bildungs- und Erziehungsauftrag der Schulen neben der Vermittlung von Fachkompetenzen auch die Prävention und Förderung gesundheitsbezogener Kompetenzen umfasst: Schülerinnen und Schüler sollen lernen, »Freude an der Bewegung und am gemeinsamen Sport zu entwickeln, sich gesund zu ernähren und gesund zu leben« (Schulgesetz NRW, 2005, Erster Teil, §2 Absatz 5). Die Befähigung der Schülerinnen und Schüler zu gesundheitsbewusstem *Handeln* (hier = Gesundheitskompetenz) wird in den Bildungsstandards der Kultusministerkonferenz im Fach Biologie für den Mittleren Schulabschluss hervorgehoben (KMK, 2005a). Weiterhin weist die Kultusministerkonferenz in ihren Empfehlungen darauf hin, dass die Schule dazu anregen und befähigen sollte, durch regelmäßige sportliche Betätigung die körperliche *und* geistige Leistungsfähigkeit bis ins hohe Alter zu erhalten (KMK, 2005b)². Die große Bedeutung, die im schulischen Kontext der Förderung gesundheitsbezogener Kompetenzen mittlerweile beigemessen wird, spiegelt sich u.a. in Bestrebungen zum Gesundheitsmanagement an Schulen wider. Das hessische Kultusministerium hat beispielsweise ein Verfahren zur Zertifizierung von Schulen als »gesundheitsfördernde Schule« eingeführt, das Ernährungs- und Verbraucherbildung, Bewegung und Wahrnehmung sowie Sucht- und Gewaltprävention als verpflichtende Teilzertifikate umfasst (Hessisches Kultusministerium, 2008). Auch die WHO weist explizit darauf hin, dass Erziehung und Gesundheit untrennbar miteinander verknüpft sind (siehe Hessisches Kultusministerium, 2008).

Gesundheitskompetenz: Das eigene Modell

Erstes Ziel des Forschungsprojekts »Gesundheitskompetenz – Modellbildung und Validierung« aus dem Schwerpunktprogramm 1293 war es, ein umfassendes, theoretisch fundiertes und empirisch abgesichertes Modell der Gesundheitskompetenz zu entwickeln. Analog zur Konzeption des Kompe-

2 Hier kann ein verwandtes Projekt *Verbesserung der Verkehrserziehung an weiterführenden und beruflichen Schulen* erwähnt werden, welches – durchgeführt im Auftrag der Bundesanstalt für Straßenwesen und in Kooperation mit den Kultusbehörden der Länder – auch im Kontext der Bildungsstandards zu sehen ist. Es geht um die Entwicklung von Leistungsstandards und deren Evaluation für den Bereich Verkehrserziehung als dritte Komponente zusammen mit Gesundheits- und Umwelterziehung. Die Entwicklung des hier eingesetzten computerapplizierten Verkehrsquiz' ging einher mit Untersuchungen der basismotorischen Kompetenzen im Zusammenhang mit dem Fahrradfahren im Fahrradparcours, motorischen Aufgaben und Selbsteinschätzungen der Kinder und Jugendlichen. Inzwischen ist von der KMK entschieden worden, dieses Verkehrsquiz bundesweit einzusetzen.

tenzbegriffs im Schwerpunktprogramm wird die Gesundheitskompetenz dabei als eine primär kognitive Kompetenz betrachtet (vgl. Weinert, 2001). Aufbauend auf diesem Kompetenzbegriff und einer umfassenden Literaturanalyse wurde zunächst eine Arbeitsdefinition entwickelt, die Gesundheitskompetenz als

... eine Sammlung von Fähigkeiten und Fertigkeiten versteht, über die jemand verfügen muss, um im Alltag und im Umgang mit dem Gesundheitssystem so handeln zu können, dass es sich positiv auf seine Gesundheit und sein Wohlbefinden auswirkt.

Um eine umfassende theoretisch fundierte Modellbildung vorzunehmen, wurde basierend auf der Literaturanalyse eine Expertenbefragung nach der Methode des Concept Mappings (Trochim, 1989; Kane & Trochim, 2007) durchgeführt. Als Experten wurden Personen aus dem Gesundheitsbereich sowie Wissenschaftlerinnen und Wissenschaftler aus der Gesundheits- und Kompetenzforschung eingeladen, an der zweistufigen Befragung teilzunehmen. In der ersten Befragung wurde anhand eines Brainstorming eine möglichst erschöpfende Sammlung von Fähigkeiten und Fertigkeiten erhoben, die mit der oben genannten Definition der Gesundheitskompetenz assoziiert wurden. Von den insgesamt N=99 teilnehmenden Expertinnen und Experten wurden 328 Aussagen produziert, die nach Bereinigung von Redundanzen auf 105 Aussagen verdichtet werden konnten.

Die aus der Aggregation der Sortierungen der 27 Teilnehmer gewonnene Datenmatrix (*group similarity matrix*), bildete den Input für das Multidimensionale Skalieren (MDS). Mit Hilfe der Concept Mapping Software *CS Global 4.0* wurde eine zweidimensionale Lösung für das MDS gerechnet (s. Kane & Trochim, 2007). Auf der Basis der x-y-Koordinaten dieser MDS-Lösung wurde in einem weiteren Schritt des Concept Mappings eine hierarchische Clusteranalyse gerechnet. Eine 9-Cluster-Lösung zeigte nach inhaltlichen Kriterien die beste Lösung für das zu entwickelnde Modell der Gesundheitskompetenz.

Die mit Global CS berechnete 2-dimensionale MDS-Lösung wies einen Stresswert von 0.23 auf, was einen relativ guten Fit der Daten, verglichen mit dem bei *Concept-Mapping*-Untersuchungen zu erwartenden mittleren Stresswert, indiziert. Der Stresswert für die mit SPSS berechnete 3-dimensionale Lösung lag bei 0.015. Dieser Wert weist auf einen sehr guten Fit der Daten, entsprechend der klassischen Kategorisierung nach Kruskal, hin. Die Clusterlösung auf der Grundlage der 3-dimensionalen MDS-Lösung erwies sich eindeutiger interpretierbar als die auf der Grundlage der 2-dimensionalen Lösung. Dies gab den Ausschlag, die 3-dimensionale MDS-Lösung als Grundlage für die weitere Modellentwicklung zu nutzen. Folgende neun Cluster konnten identifiziert werden (siehe Abb. 3):

- (1) die Fähigkeit zu Selbstregulation und Selbstdisziplin;
- (2) die Fähigkeit zur Wahrnehmung der eigenen Bedürfnisse und Gefühle sowie ein hohes Körperbewusstsein;
- (3) die Bereitschaft zur Verantwortungsübernahme für die eigene Gesundheit;
- (4) gesundheitsbezogene Grundfertigkeiten, insbesondere die Fähigkeit, gesundheitsrelevante Texte lesen und verstehen zu können (*literacy*) und gesundheitsrelevante mathematische Aufgabenstellungen lösen zu können (*numeracy*);
- (5) die Fähigkeit, gesundheitsrelevante Informationen angemessen interpretieren und nutzen zu können, wozu auch ein bestimmtes Maß an medizinisch-biologischem Grundwissen nötig ist;
- (6) die Fähigkeit, sich gesundheitsrelevante Informationen beschaffen zu können;
- (7) die Fähigkeit, innerhalb des Gesundheitssystems navigieren und handeln zu können, sowie über das dazu notwendige Systemwissen zu verfügen;
- (8) die Fähigkeit zur Kommunikation und Kooperation bezüglich gesundheitsrelevanter Inhalte;
- (9) förderliche Persönlichkeitseigenschaften.

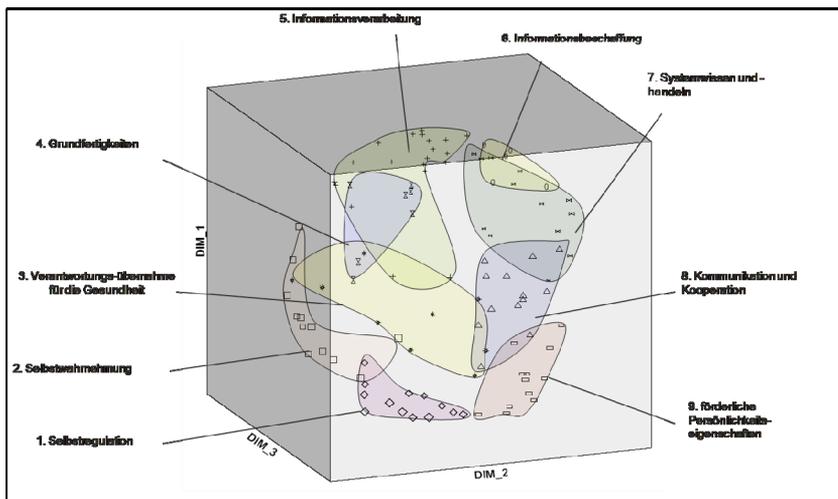


Abbildung 3: Ergebnis des 3-dimensionalen Concept Mapping mit der 9-Cluster-Lösung

Dem dreidimensionalen Concept Map lassen sich auch Informationen über die Beziehungen der Teilkomponenten zueinander entnehmen. So können die Cluster Informationsbeschaffung und -verarbeitung sowie Systemwissen und -handeln als systembezogene Komponenten des Modells und die ihnen ge-

nüberliegenden Cluster Selbstregulation, Selbstwahrnehmung und Verantwortungsübernahme links unten im Würfel als personenbezogene Komponenten interpretiert werden. Dazwischen liegen in der Mitte die Cluster der Grundfertigkeiten und Kommunikation und Kooperation, welche plausibel als zwischen der Person und dem System vermittelnde Variablen gedeutet werden können.

Als Ergebnis dieses hier nur in Stichworten beschriebenen Prozesses konnte ein erstes Strukturmodell der Gesundheitskompetenz formuliert werden, das folgende 8 Komponenten beinhaltet (siehe Abb. 4):

- (1) die Fähigkeit zur Wahrnehmung der eigenen Bedürfnisse und Gefühle sowie eine hohes Körperbewusstsein
- (2) die Fähigkeit zu Selbstregulation und Selbstdisziplin
- (3) die Fähigkeit zur Verantwortungsübernahme für die eigene Gesundheit bzw. die Fähigkeit, Gesundheit als ein aktiv herzustellendes Ziel verstehen zu können
- (4) die Fähigkeit, sich gesundheitsrelevante Informationen beschaffen zu können
- (5) die Fähigkeit ,diese angemessen verstehen und interpretieren zu können
- (6) die Fähigkeit, innerhalb des Gesundheitssystems handeln zu können und über das dazu notwendige Systemwissen zu verfügen
- (7) gesundheitsbezogene Grundfertigkeiten, insbesondere die Fähigkeit, gesundheitsrelevante Texte lesen und verstehen zu können (literacy) sowie gesundheitsrelevante mathematische Grundfertigkeiten (numeracy)
- (8) die Fähigkeit zur Kommunikation und Kooperation über gesundheitsrelevante Inhalte

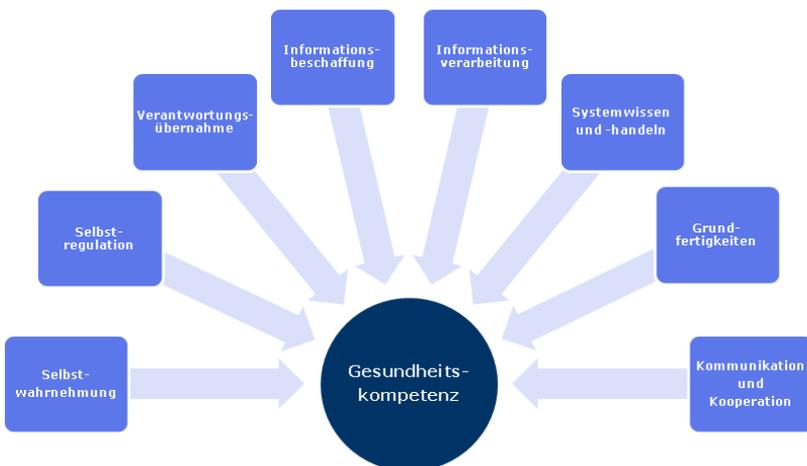


Abbildung 4: Strukturmodell der Gesundheitskompetenz

Auf der Grundlage des *Concept Maps* sowie weiterer strukturierender theoretischer Überlegungen wird nun Gesundheitskompetenz als ein Netz gesehen, bestehend aus

- Grundlegenden Fertigkeiten (literacy/numeracy)
- Handlungskompetenz: Sie kann in die vier Kompetenzbereiche unterteilt werden (a) Navigieren und Handeln im Gesundheitssystem, (b) Kommunikation und Kooperation, (c) Informationsbeschaffung und -verarbeitung sowie (d) Selbstwahrnehmung und Selbstregulation
- Wissen: Die Wissenskomponente wird durch System- und Gesundheitswissen repräsentiert, wobei Systemwissen im Concept Map in Cluster (7) und Gesundheitswissen in Cluster (5) enthalten ist
- Motivation: Cluster (3), Bereitschaft zur Verantwortungsübernahme für die eigene Gesundheit, bildet den motivationalen Teil der Gesundheitskompetenz

Persönlichkeitseigenschaften sind im Sinne des Modells keine Strukturkomponenten, werden jedoch in weiteren Untersuchungen in ihrer Beziehung zum Modell zu analysieren sein.

Die Komponenten des Modells sind kognitiv konzipiert, auch wenn einzelne Strukturkomponenten, insbesondere die motivationale Komponente zur Verantwortungsübernahme für die eigene Gesundheit, den Rahmen eines engen Kognitionsbegriffs verlassen. Das Strukturmodell stellt eine Erweiterung bereits existierender Modelle der Gesundheitskompetenz dar, insofern es vormals getrennt betrachtete Komponenten wie Grundfertigkeiten, Kommunikation, Informationsverarbeitung und Handlungsbereitschaft (im Sinne der Verantwortungsübernahme für die Gesundheit) integriert und miteinander in Beziehung setzt (vgl. Nutbeam 2000; Schulz/Nakamoto 2005; Kriegesmann u.a. 2005; vgl. auch Soellner u.a. 2009). Eine derart ganzheitliche Betrachtung des Konzeptes unter Einbezug verschiedenster notwendiger Strukturkomponenten wurde bislang noch nicht geleistet. Im Vergleich mit bisherigen Modellen der Gesundheitskompetenz weist das in diesem Forschungsprojekt erarbeitete Modell wesentliche Besonderheiten auf:

- (1) Das Modell wurde umfassend und systematisch unter Einbezug von Expert(inn)en aus dem Gesundheitsbereich und der Kompetenzforschung konzipiert.
- (2) Das Konstrukt der Gesundheitskompetenz wurde auf Basis eines Kompetenzbegriffs entwickelt, welcher dem Modell einen klaren theoretischen Rahmen gibt.
- (3) Es wurde ein Gesamtmodell der Gesundheitskompetenz geschaffen, welches über die heterogene Vielfalt gesundheitsrelevanter Situationen anwendbar sein soll, aber dennoch situations- und kontextspezifisch ausformuliert werden kann (i. S. eines Wechselspiels zwischen den Strukturkomponenten und den jeweiligen situations- und kontextspezifischen Anforderungen).

- (4) Das vorgelegte Modell erweitert bisherige Modelle der Gesundheitskompetenz in seiner Gesamtheit und weist diese als Teilbereiche der Gesundheitskompetenz aus.
- (5) Dabei ergänzt es bisherige Modelle, v.a. um die Kompetenzbereiche der Selbstregulation und Selbstwahrnehmung sowie um den motivationalen Aspekt zur aktiven Verantwortungsübernahme für die eigene Gesundheit.

Literatur

- ABEL, T. & Bruhin, E. (2003). Health Literacy / Wissensbasierte Gesundheitskompetenz (4. Aufl.). In Bundeszentrale für gesundheitliche Aufklärung (Hrsg.), *Leitbegriffe der Gesundheitsförderung* (S. 128-131). Schwabenheim a. d. Selz: Verlag Peter Sabo.
- ABEL, T. & Walter, E. (2002). Individuelles Verhalten und soziale Verantwortung für Gesundheit: Anmerkungen zur Wissenschaft und Praxis der Gesundheitsförderung. In P. Heusser (Hrsg.), *Gesundheitsförderung – Eine neue Zeitforderung* (S. 31-42). Frankfurt/M.: Peter Lang AG.
- AMERICAN Medical Association, Ad Hoc Committee on Health Literacy. (1999). Health literacy: Report of the Council on Scientific Affairs. *JAMA*, 281, 552-557. Zugriff am 21.01.2009 von <http://www.cmaj.ca/cgi/ijlink?linkType=ABST&journalCode=jama&resid=281/6/552>.
- ANDRUS, M. R. & Roth, M. T. (2002). Health Literacy: A Review. *Pharmacotherapy*, 22 (3), 282-302.
- BAKER, D. W. (2006). The meaning and the measure of health literacy. *Journal of General Internal Medicine*, 21 (8), 878-883.
- DESECO (2005). *Definition und Auswahl von Schlüsselkompetenzen*. Zugriff am 07.09.2008 unter <http://www.deseco.admin.ch/bfs/deseco/en/index/03/04.parsys.97111.downloadList.89603.DownloadFile.tmp/2005.dskcexecutivesummary.ge.pdf>
- EUROPEAN Patients Forum (2008). *EPF-Conference 2008*. Zugriff am 21.01.2009 unter http://www.eu-patient.eu/policy/attached_documents/EPF_HealthLiteracyConference_2008_Key_Conclusions.pdf
- HESSISCHES Kultusministerium (2008). *Schule & Gesundheit Hessen 2002 – 2011: Grundlagen – Strategien – Meilensteine*. Zugriff am 18.01.2009 unter http://www.schuleundgesundheit.hessen.de/fileadmin/content/Konzept/Grundlagenpapier_2002_2011.pdf
- ISHIKAWA, H., Nomura, K., Sato, M. & Yano, E. (2008). Developing a measure of communicative and critical health literacy: A pilot study of Japanese office workers. *Health promotion international*, 23, 269-274.
- KANE, M. & Trochim, W. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks: Sage.

- KICKBUSCH, I. (2006). Gesundheitskompetenz. *News: Public Health Schweiz*, Nr. 3-2006, S. 10. Zugriff am 14.03.2008 unter http://www.public-health.ch/client/publichealth/file/PHS_Newsletter_2006_3.pdf
- KICKBUSCH, I., Maag, D. & Saan, H. (2005). *Enabling healthy choices in modern health societies*. Paper presented at the European Health Forum 2005. Zugriff am 19.01.2009 unter http://www.ilonakickbusch.com/health-literacy/Gastein_2005.pdf
- KMK (2005a) (Hrsg.). *Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.
- KMK (2005b). *Die Bedeutung des Schulsports für lebenslanges Sporttreiben. Gemeinsame Erklärung der Präsidentin der Kultusministerkonferenz, des Präsidenten des Deutschen Sportbundes und des Vorsitzenden der Sportministerkonferenz*. Karlsruhe, 12. Dezember 2005. Zugriff am 08.09.2009 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2005/2005_12_12-Schulsport-lebenslanges-Sporttreiben.pdf
- KRIEGESMANN, B., Kottmann, M., Masurek, L. & Nowak, U. (2005). Kompetenz für eine nachhaltige Beschäftigungsfähigkeit. (*Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Forschungsbericht, Fb 1038*). Bremerhaven: Wirtschaftsverlag NW Verlag für neue Wissenschaft GmbH.
- LÖSEL, F. & Bliesener, T. (2003). *Aggression und Delinquenz unter Jugendlichen*. München: Luchterhand.
- MAAG, D. (2007). *Gesundheitskompetenz bezüglich Ernährung, Bewegung und Gewicht*. Unveröffentlichte Dissertation, Università della Svizzera italiana, Lugano.
- NUTBEAM, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health Promotion International*, 15 (3), 259-267.
- PARKER, R. M., Baker, D. W., Williams, M. V. & Nurss, J. (1995). The Test of Functional Health Literacy in Adults: A new instrument for measuring patients' literacy skills. *Journal of General Internal Medicine*, 10, 537-541.
- PLEASANT, A. & Kuruvilla, S. (2008). A tale of two health literacies: Public health and clinical approaches to health literacy. *Health Promotion International*, 23 (2), 152-159.
- RATZAN, S. C. & Parker, R. M. (2000). Introduction. In C.R. Selden, M. Zorn & S. C. Parker (Eds.), *Current Bibliographies in Medicine: Health Literacy* (CBM 2000-1). Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services.
- RUDD, R., Kirsch, I. & Yamamoto, K. (2004). *Literacy and health in America*. Princeton, NJ: Harvard School of Public Health & Center for Global Assessment.
- SCHULGESETZ des Landes Nordrhein-Westfalen in der Fassung vom 15.02.2005. (Erster Teil §2 Absatz 5). Zugriff am 08.09.2009 unter http://www.schulministerium.nrw.de/BP/Schulrecht/Gesetze/SchulG_Info/Schulgesetz.pdf

- SCHULZ, P. & Nakamoto, K. (2005). Emerging themes in health literacy. *Studies in Communication Sciences*, 5 (2), 1-10.
- SCHWEIZERISCHE Gesellschaft für Gesundheitspolitik. (2003). *The Future Patient in Switzerland: Das Gesundheitssystem der Zukunft aus der Sicht von Bürgerin und Bürger. (SGGP-Schriftenreihe, Band 71)*. Zürich: SGGP.
- SOELLNER, R., Huber, S., Lenartz, N. & Rudinger, G. (ang.). Gesundheitskompetenz – ein vielschichtiger Begriff. *Zeitschrift für Gesundheitspsychologie*.
- SPLIETH, C., Müller, M., Kohlmann, T., John, U. & Hoffmann, W. (2007). *GeKo^{KiDS}: Evaluation eines bevölkerungsbezogenen Programms zur Stärkung der gesundheitsbezogenen Kompetenzen für die Jahrgangsstufe 5/6 – Design und Implementation*. Zugriff am 21.01.2009 unter http://www.medizin.uni-greifswald.de/icm/abt_vcm/projekte/GeKoKidS_Poster_8_07.pdf
- STAUDT, E., Kailer, N., Kriegesmann, B., Meier, A. J., Stephan, H. & Ziegler, A. (2002). Kompetenz und Innovation: Eine Bestandsaufnahme jenseits von Personalmanagement und Wissensmanagement. In E. Staudt (Hrsg.), *Kompetenzentwicklung und Innovation: Die Rolle der Kompetenz bei Organisations-, Unternehmens- und Regionalentwicklung* (S. 127–235). Münster: Waxmann.
- STECKELBERG, A., Hülfenhaus, C., Kasper, J., Rost, J. & Mühlhauser, I. (2007). How to measure critical health competences: Development and validation of the Critical Health Competence Test (CHC Test). *Advances in Health Sciences Education*, online first. Zugriff am 15.01.2009 unter <http://www.springerlink.com/content/6761364581252861/?p=b1485d1d8166479195a6262451072aa7&pi=0>.
- TIPPELT, R., Mandl, H. & Straska, G. (2003). Entwicklung und Erfassung von Kompetenz in der Wissensgesellschaft – Bildungs- und wissenstheoretische Perspektiven. In I. Gogolin und R. Tippelt (Hrsg.), *Innovation durch Bildung. Beiträge zum 18. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft* (S. 349-369). Opladen: Leske+Budrich.
- TONES, K. (2002). Health literacy: New wine in old bottles? *Health Education Research*, 17 (3), 287-290.
- TROCHIM, W. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, 12, 1-16.
- WANG, J. & Schmid, M. (2007). *Regionale Unterschiede in der Gesundheitskompetenz in der Schweiz*. Zürich: Institut für Sozial- und Präventivmedizin. Zugriff am 21.09.2009 unter http://www.ispmz.ch/downloads/HintergrundGesundheitskompetenz_de_070607.pdf
- WEINERT, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and Selecting Key Competencies* (pp. 45-65). Göttingen: Hogrefe.
- WHO (1998). *Health promotion glossary*. Zugriff am 21.01.2009 unter http://www.who.int/hpr/NPH/docs/hp_glossary_en.pdf
- WILSON, M. (2005). *Constructing measures: An item-response approach*. Mahwah, N.J.: Lawrence Erlbaum.

ZARCADOOLAS, C., Pleasant, A. F. & Greer, D. S. (2006). *Advancing health literacy: A framework for understanding and action*. San Francisco, CA: Jossey-Bass.

Teil III

Umfragen

Das ZEM erstellt für Auftraggeber aus Wissenschaft und Wirtschaft regelmäßig durch Umfragen Studien und Datensätze. So wurden alleine in den letzten drei Jahren mehr als 200 Erhebungen vom ZEM durchgeführt. In der Einleitung sind einige der Partner genannt, im Teil V werden einige dieser Studien kurz referiert. Hier im Teil III werden vier Beispiele solcher Studien gegeben, die bei aller thematischer Unterschiedlichkeit (*Rektorenranking, Unternehmenssteuerung mit dem Net Promoter Score, Bedarfsanalyse für Gründerzentren, Lebensverlängernde Maßnahmen*)¹ die Gemeinsamkeit der Kontaktierung schwieriger Zielgruppen mit sehr spezifischen Fragestellungen haben. Ergänzt werden diese Studien durch einen externen Beitrag zum *Net Promoter Score*.

1 *Kursive Passagen* benennen bzw. umschreiben die Themen der Artikel

Thomas Krüger & Georg Rudinger Rektoren-Ranking: Rektor des Jahres 2009¹

In einer Umfrage des Deutschen Hochschulverbandes haben über 3.000 Wissenschaftlerinnen und Wissenschaftler Anforderungsprofile für den Rektoren bzw. Präsidenten² ihrer Hochschule erstellt und die jeweiligen Amtsinhaber danach bewertet. Über welche Kompetenzen sollte aus Sicht der Professorenschaft das ideale akademische Oberhaupt verfügen? Und welcher Rektor entspricht am besten den Wünschen und Erwartungen »seiner« Wissenschaftler? Ergebnisse des ersten Rektoren-Rankings in Deutschland.

Die Leitung einer Hochschule ist eine anspruchsvolle Managementaufgabe. Es gilt Interessen verschiedener Gruppen – neudeutsch »Stakeholder« – zu berücksichtigen und zu integrieren; konkret der Politik (Ministerium), der Studentenschaft und deren zukünftigen Arbeitgebern, aber natürlich auch der Kollegen und Mitarbeiter der Hochschulen. Die wachsende Autonomie der Hochschulen, der zunehmende Wettbewerb bis hin zu Exzellenzinitiativen, aber auch die Einführung von »Aufsichtsräten« stellen neue Forderungen an Rektoren.

Zur Messung der Qualität der Hochschulleitung lassen sich verschiedene mehr oder minder sinnvolle Parameter definieren. So ist es sicherlich auch eine Auszeichnung für die Leitung, wenn eine Hochschule als exzellent eingestuft wird oder auf Spitzenplätzen bei Bewerberzahlen landet. Die hier vorgestellte Studie nimmt eine andere Perspektive ein, nämlich die der Wissenschaftler, die an den Hochschulen forschend und lehrend tätig sind. Im Auftrage des Deutschen Hochschulverbandes untersuchte das Zentrum für Evaluation und Methoden, Universität Bonn, welche Erwartungen und Anforderungen die Wissenschaftler an ihre Hochschulleitung stellen – und welcher Rektor diesen am besten gerecht wird.

Die Mitglieder des deutschen Hochschulverbandes (DHV), von denen dem DHV eine E-Mail-Adresse vorliegt, wurden elektronisch kontaktiert und gebeten, im Rahmen einer Online-Befragung aus einer Liste die Eigenschaften auszuwählen, die der Rektor ihrer Hochschule idealerweise erfüllen soll-

-
- 1 Der Beitrag ist ein Nachdruck aus »Forschung & Lehre« (4/2009) Wir danken dem Deutschen Hochschulverband für die freundliche Erlaubnis des Nachdrucks.
 - 2 Im Folgenden wird der besseren Lesbarkeit halber einheitlich von Rektor bzw. Rektoren gesprochen, auch wenn an einigen Hochschulen die Bezeichnung Präsident verwendet wird. Maskuline Personen- und Funktionsbeschreibungen gelten grundsätzlich immer für Männer und Frauen.

te. Die Eigenschaften wurden einer Einzelfallstudie³ entnommen und zusammen mit Experten erweitert und verfeinert. Zudem wurde erfragt, wie stark die jeweiligen Eigenschaften beim aktuellen Rektor ausgeprägt sind. Der Rektor, welcher in seinem »Ist-Profil« im Urteil seiner Universitätsmitglieder am besten den mit seinem Amt verbundenen Anforderungen gerecht wird, ist »Rektor des Jahres 2009«. Diese Auszeichnung des Hochschulverbandes soll in den nächsten Jahren regelmäßig verliehen werden. Die durchaus positive Resonanz auf dieses Rektorenranking umfasst Anregungen, dies u.a. auf die Kanzler, auf die Verwaltung etc. auszudehnen.

Statistische Überlegungen und andere Erwägungen gingen der Umfrage voraus: Um für die einzelnen Rektoren eine hinreichend große Anzahl von Bewertungen erreichen zu können, wurden von vornherein bei der Online-Umfrage nur größere Hochschulen mit mindestens 1.000 Studenten berücksichtigt. Die Umfrage beschränkte sich entsprechend auf Universitäten und ihnen gleichgestellte Hochschulen, da der DHV an den anderen Hochschulen nicht so stark vertreten ist⁴. Damit überhaupt eine sinnvolle Bewertung möglich ist, sollte der Rektor zum Start der Befragung schon mindestens 100 Tage im Amt sein. Andererseits sollte seine Amtszeit nicht im zeitlichen Umfeld (von +/- einem Monat) der Verleihung des Preises »Rektor des Jahres« am 30. März enden. Nach diesen Kriterien wurden 89 Hochschulen in die Befragung eingeschlossen⁵, deren Wissenschaftler online ihren Rektor bewerten konnten. Um auszuschließen, dass einzelne Personen durch mehrfaches Antworten die Ergebnisse verfälschten, konnte der Online-Fragebogen jeweils nur einmal ausgefüllt werden. Allen Mitgliedern, von denen dem Hochschulverband eine E-Mail-Adresse vorlag, wurde dazu ein eigener Log-in zugesendet. Sobald mit einem Log-in eine Bewertung abgegeben worden war, wurde dieser Log-in gesperrt.

Die Einladungen mit den Log-ins wurden vom Hochschulverband zwischen dem 6. und 9. Januar per E-Mail an 16.429 Mitglieder verschickt. Am 20. Januar wurden diese Mitglieder nochmals per E-Mail angeschrieben und zur Teilnahme eingeladen. Am 6. Februar wurde die Datenerhebung beendet. Insgesamt haben 4.516 Wissenschaftler auf die Einladung reagiert und den Fragebogen aufgerufen. In dem Online-Fragebogen musste zunächst die jeweilige Hochschule angegeben werden. Wenn die Hochschule nicht den oben genannten Kriterien entsprach, wurde die Befragung mit einem Hinweis

3 Braunecker, C. (2006). *Anforderungsprofil für den neuen Rektor der Universität Wien*. Zugriff am 09.09.2009 unter

http://senat.univie.ac.at/fileadmin/user_upload/senat/aktuelles/Rektorbericht.pdf

4 In anstehenden Rankings sollen auch Rektoren kleinerer Universitäten und Fachhochschulen berücksichtigt werden.

5 Namen, Bezeichnungen und Amtszeiten der Rektoren wurden durch das ZEM telefonisch bei den jeweiligen Pressestellen erfragt. Studentenzahlen sind Angaben der KMK.

beendet und der Log-in gesperrt. Knapp 400 Personen konnten daher nicht an der Befragung teilnehmen. Gut 500 weitere Personen haben den Fragebogen nicht zu Ende ausgefüllt. Somit verbleiben 3.619 Personen. Für das Ranking sollten aus statistischen Erwägungen nur Rektoren berücksichtigt werden, für die aus ihrer Hochschule mindestens dreißig Bewertungen vorlagen⁶. Dies hat die Zahl der Hochschulen noch einmal – nämlich auf 52 – reduziert. Tabelle 1 listet die 52 Hochschulen auf, welche nach diesem zweistufigen Verfahren in die Endauswertung kamen. Das entspricht den Antworten von 3.116 Wissenschaftlern.

Den Befragungsteilnehmern aus den 89 Hochschulen wurde eine Liste von 18 Eigenschaften in zufälliger Reihenfolge vorgelegt. Daraus sollten die drei bis fünf wichtigsten Eigenschaften ausgewählt werden, über die der Rektor der jeweiligen Hochschule verfügen sollte. In Abbildung 1 sind die Eigenschaften dargestellt, geordnet nach der von den Wissenschaftlern bewerteten Wichtigkeit. Führungskompetenz, Respekt für die Unterschiedlichkeit der Wissenschaftskulturen und Kenntnis des Universitätsbetriebes sind demnach die drei wichtigsten Eigenschaften eines Rektors, die von jeweils über 40% der Teilnehmer ausgewählt wurden. Diese Eigenschaften werden aber nicht an allen Hochschulen gleichermaßen für wichtig erachtet: Zum Beispiel wurde »Respekt für die Unterschiedlichkeit der Wissenschaftskulturen« nur von 16% der Wissenschaftlern der Universität Hohenheim, aber von 67% der Universität Bielefeld als eine der wichtigen Eigenschaften des Rektors gewählt. Dies genau ist das Charakteristikum unserer Befragung: Die Rektoren werden gemäß der für ihre Hochschule als wichtig erachteten Kompetenzen bewertet. Dass es diesbezüglich große Unterschiede gibt veranschaulicht Abbildung 1 in den linken und rechten Profilen, welche die Extremgewichten je einzelner Universitäten darstellen.

6 Dass dabei auch gut positionierte Rektoren, z.B. der Rektor der Universität Wuppertal, mit 29 Antworten, nicht berücksichtigt werden konnten, ist bedauerlich. Wir hoffen daher, dass bei der nächsten Welle die Teilnahmequote noch höher ist!

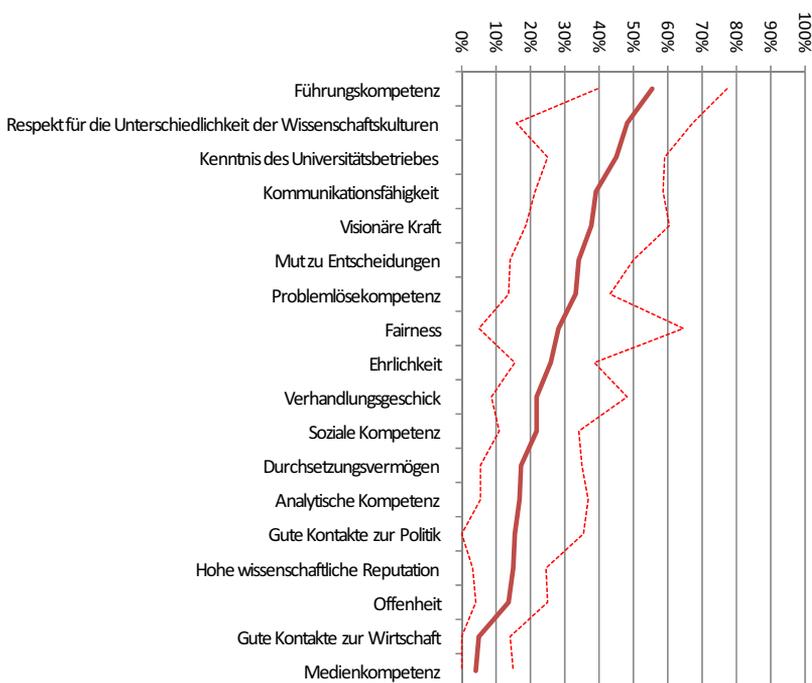


Abbildung 1: Wichtigkeit der Eigenschaften für das Amt eines Rektors. Wahl unter die wichtigsten Eigenschaften: Mittlerer Anteil über alle Hochschulen (Mitte) und kleinster (links) bzw. größter (rechts) Anteil auf Ebene einzelner Hochschulen.

Gute Kontakte zu Politik und zu Wirtschaft sind nach Meinung der meisten Befragten eher unwichtig für einen Rektor. Die für Manager aus der Wirtschaft sehr wichtige Medienkompetenz landet bei der Umfrage auf dem letzten Platz. Bemerkenswert ist aber auch, dass die Wissenschaftler von ihrem Rektor nicht unbedingt erwarten, dass dieser eine hohe wissenschaftliche Reputation hat.

Nach Bestimmung der drei bis fünf wichtigsten Eigenschaften sollten die Teilnehmer ihren aktuellen Rektor bezüglich aller 18 Eigenschaften jeweils mit Schulnoten von 1 (sehr gut) bis 6 (ungenügend) bewerten. Sie wurden den Wissenschaftlern wieder in zufälliger Reihenfolge dargeboten. Wenn ein Urteil nicht möglich oder gewollt war, konnten Eigenschaften übersprungen werden.

In Abbildung 2 sind die Mittelwerte der Noten der Rektoren in den Eigenschaften über alle Befragungsteilnehmer, d.h. eben auch über alle Hochschulen abgetragen, sowie die jeweils besten (linkes Profil) bzw. schlechtesten (rechtes Profil) Eigenschafts-Noten der Rektoren einzelner Universitäten. Am besten wird über alle Rektoren hinweg ihre Kenntnis des Universitätsbetrie-

bes bewertet (Durchschnittsnote: 2,6). Hier schneidet kein einziger Rektor schlechter als mit einer 4,2 ab. Und die beste Bewertung ist hier ein 1,6 – für den Rektor der MH Hannover. Mankos werden bei »weichen« Faktoren wie »Kommunikationsfähigkeit« und »Soziale Kompetenz« gesehen. Hier erhalten die am schlechtesten bewerteten Rektoren eine 5,2 bzw. 5,1. Auch die Besten erreichen hier »nur« 2,2 (Rektor Universität Konstanz bzw. Ruhr-Universität Bochum). Bezüglich der »Visionären Kraft« liegt die Durchschnittsnote bei einer eher bescheidenen 3,4. Auch hier bewerten die Wissenschaftler der Ruhr-Universität Bochum ihren Rektor mit einer Note von 1,8 am besten.

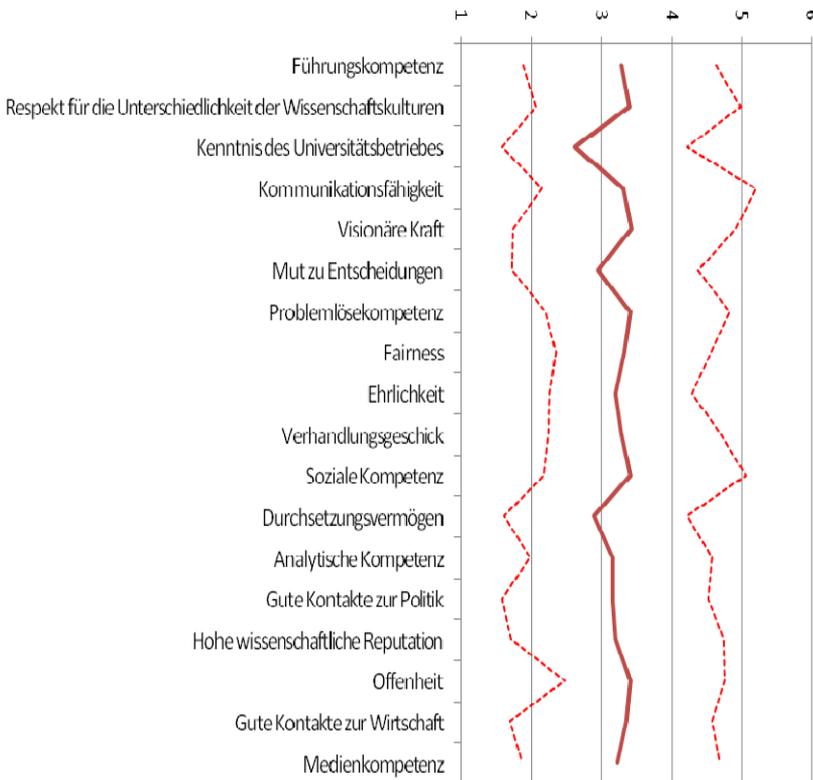


Abbildung 2: Benotung der Rektoren

Eigenschaftsnoten: Mittlere Benotung über alle Teilnehmer (Mitte) und die besten (links) bzw. schlechtesten (rechts) mittlere Noten auf Ebene einzelner Rektoren. Von 1=sehr gut bis 6=ungenügend

Aus den 18 Noten zu den Eigenschaften haben wir pro Hochschule eine gewichtete Eigenschaftsnote berechnet. Gewichtet wurde hier nach der von

den Wissenschaftlern der jeweiligen Hochschule gesehene Bedeutung der Eigenschaften für einen idealen Rektor. Nur drei Rektoren kamen hier auf einen Durchschnitt besser als 2,5: Dies waren die Rektoren der Ruhr-Universität Bochum (2,1), der Medizinische Hochschule Hannover (2,2) und der Universität Konstanz (2,4). Über alle Teilnehmer lag die Durchschnittsnote bei 3,2, wobei vier Rektoren schlechter als 4,0 abschnitten.

Am Ende der Befragung wurden die Teilnehmer schließlich gebeten, mit einer Note – wieder zwischen 1 und 6 – die Leistungen ihres Rektors zusammenfassend zu bewerten. Über alle Teilnehmer gemittelt erhalten die deutschen Rektoren als Durchschnittsnote ein Befriedigend (2,9). Aber auch hier finden sich deutliche Unterschiede. Primus ist mit einer 1,7 wieder der Rektor der Ruhr-Universität Bochum, während der Ultimus mit einer 4,3 nur gerade noch ausreichend bewertet wird.

Für das Ranking wurde als Gesamtnote das Mittel aus dieser Leistungsnote und der gewichteten Eigenschaftsnote bestimmt. In Tabelle 1 sind die (von den 89) verbliebenen 52 Hochschulen mit mindestens 30 Bewertungen alphabetisch nach Standort aufgelistet. Für die Top 10 der Rektoren ist auch die Gesamtnote und der Rangplatz angegeben. Der erste Platz geht an Prof. Dr. Elmar Weiler, Rektor der Ruhr-Universität Bochum. Ihm folgen Prof. Dr. Arnold Ganser, Präsident der MH Hannover und Prof. Dr. Dr. h.c. Gerhart von Graevenitz, Rektor der Universität Konstanz.

Tabelle 1: Berücksichtigte Hochschulen, Rang (Top 10), Gesamtnote (Top 10) und Anzahl Bewertungen (N).

<i>Hochschule</i>	<i>Rang</i>	<i>Gesamtnote</i>	<i>N</i>
RWTH Aachen	9	2,42	60
Universität Augsburg			38
Otto-Friedrich-Universität Bamberg			31
Universität Bayreuth			55
Freie Universität Berlin			91
Humboldt-Universität zu Berlin			53
Technische Universität Berlin			52
Universität Bielefeld	7	2,37	52
Ruhr-Universität Bochum	1	1,93	101
Technische Universität Braunschweig			60
Universität Bremen			55

Technische Universität Darmstadt			63
Technische Universität Dortmund	4	2,31	40
Technische Universität Dresden			52
Universität Duisburg-Essen			76
Friedrich-Alexander-Universität Erlangen-Nürnberg			102
Albert-Ludwigs-Universität Freiburg			55
Justus-Liebig-Universität Gießen			71
Georg-August-Universität Göttingen			111
Ernst-Moritz-Arndt-Universität Greifswald			33
Martin-Luther-Universität Halle-Wittenberg			53
Universität Hamburg			107
Medizinische Hochschule Hannover	2	1,98	48
Gottfried Wilhelm Leibniz Universität Hannover			57
Ruprecht-Karls-Universität Heidelberg			93
Universität Hohenheim			31
Friedrich-Schiller-Universität Jena	5	2,32	55
Universität Karlsruhe (TH)			49
Universität Kassel			50
Christian-Albrechts-Universität zu Kiel			54
Universität zu Köln	6	2,35	98
Universität Konstanz	3	2,18	41
Universität Leipzig			48
Otto-von-Guericke-Universität Magdeburg			30
Johannes Gutenberg-Universität Mainz			101
Universität Mannheim			35
Philipps-Universität Marburg			57
Technische Universität München	8	2,40	103
Ludwig-Maximilians-Universität München			103

Universität der Bundeswehr München			35
Westfälische Wilhelms-Universität Münster			91
Universität Osnabrück			36
Universität Paderborn			38
Universität Potsdam			37
Universität Regensburg			42
Universität des Saarlandes			58
Universität Siegen			46
Universität Stuttgart			60
Universität Trier			38
Eberhard Karls Universität Tübingen			71
Universität Ulm			43
Julius-Maximilians-Universität Würzburg			57

Nach längerer Diskussion haben wir uns in Übereinstimmung mit dem DHV dazu entschlossen, keine öffentlichen Aussagen zu den Plätzen 11 bis 52 zu machen. Ausschlag gab hier die Anmerkung eines Teilnehmers, dass das Rektoren-Ranking wohl auch in Ministerien kritisch beäugt werden könnte. Somit müssten gerade die Hochschulen, die schon mit einem schlechten Rektor gestraft seien, zusätzlich noch mit Nachteilen von politischer Seite rechnen. Gerne liefern die Autoren aber den Rektoren selber eine genauere Rückmeldung über ihre Position und die verschiedenen Bewertungen durch ihre Professorenenschaft.

Thomas Krüger & Astrid Mayerböck

Der Net Promoter Score als Kennziffer zur Unternehmenssteuerung

Am Zentrum für Evaluation und Methoden beschäftigen wir uns seit 2005 intensiv mit dem Net Promoter Score (NPS)¹. Dabei befassen wir uns sowohl mit der methodischen Beratung, der Durchführung und Koordination des Feldes als auch mit der statistischen Auswertung und der Interpretation der Ergebnisse. Im Jahre 2007 haben wir an der Universität Bonn einen Anwenderkongress zum NPS veranstaltet und dort auch einige Ergebnisse unserer Forschung berichtet (s. z.B. Borsch zum Net Promoter Score in diesem Band). Die Zweitautorin dieses Artikels konnte zudem im Rahmen eines Jobrotationprogramms der Universität Bonn für zwei Jahre »die Seiten wechseln« und in der Marktforschungsabteilung eines internationalen Unternehmens unter anderem NPS Studien koordinieren. Im vorliegenden Artikel möchten wir von verschiedenen Problemen im Kontext der Arbeit mit dem NPS berichten, die sich uns in der Praxis stellten und Lösungsansätze aufzeigen. Der Artikel widmet sich insbesondere den im Folgenden beschriebenen Aspekten hinsichtlich der Arbeit mit dem NPS:

– Vergleich mit Konkurrenten

Ein einzelner NPS-Wert an sich ist relativ wenig aussagekräftig. Zwar wird man sich sicherlich freuen, wenn ein Unternehmen nach Befragung der eigenen Kunden einen positiven NPS-Wert hat. Wenn jedoch Konkurrenzunternehmen einen höheren Wert verzeichnen, besteht dennoch die Gefahr, Kunden zu verlieren. Eine aussagekräftige Bewertung der eigenen Position im Markt setzt daher stets auch die Berücksichtigung von Konkurrenten voraus. Gerade wenn mehrere Konkurrenten vorhanden sind, ist es wichtig, einen mittleren NPS des Marktes zu bestimmen, um daran die eigene Position messen zu können. Dabei reicht es aber nicht, den mittleren NPS-Wert über alle Konkurrenten zu berechnen. Vielmehr müssen die Marktanteile der Konkurrenzunternehmen, also typischerweise der Kundenanteile, mitberücksich-

1 Reichheld entwickelte, gemeinsam mit Bain & Company, 2003 den so genannten Net Promoter Score (NPS), der als singuläres Maß zur Vorhersage der Unternehmensentwicklung genügen sollte. Zugrunde lagen Reichhelds Kritik an traditioneller Kundenzufriedenheitsforschung sowie Untersuchungen zum Zusammenhang zwischen Unternehmenswachstum und einer einfachen Frage: der Wahrscheinlichkeit (eingestuft auf einer 11-stufigen Skala), dass Kunden eines Unternehmens dieses an Freunde und Bekannte weiterempfehlen würden (Reichheld 2003).

tigt werden. Auf einen bedeutsamen Vergleich des eigenen Unternehmens mit dem Marktdurchschnitt gehen wir unten ein.

– Vergleich über die Zeit

Meist ist man an einer Entwicklung des NPS-Wertes über die Zeit interessiert, um festzustellen, ob dieser sich – speziell nach bestimmten Interventionen – verbessert hat – oder im Gegenteil sogar verschlechtert. Hierbei gilt es sicherzustellen, dass die beobachteten Veränderungen des Wertes tatsächlich bedeutsam, also statistisch signifikant sind. Eine Signifikanzbestimmung mit gängigen Methoden ist aufgrund der Struktur des NPS-Wertes, der als Differenz berechnet wird², nicht ohne Weiteres möglich. In unserer Forschungspraxis nutzen wir daher zur Bestimmung der Bedeutsamkeit von Veränderungen von NPS-Werten einen Simulationsansatz, den wir unten darstellen.

– Vergleich der Entwicklung in verschiedenen Bereichen

Bei dem Vergleich unterschiedlicher Einheiten, wie z.B. verschiedener Kundensegmente, Regionen oder auch Ländereinheiten, ergeben sich weitere Probleme. Hier ist ein direkter Vergleich aufgrund unterschiedlicher Ausgangssituationen meist nicht möglich. Auch für die Bewertung der zeitlichen Entwicklung muss die Ausgangssituation angemessen berücksichtigt werden. Vorschläge, wie in solchen komplexen Situationen (mehrere Einheiten, mehrere Messzeitpunkte) vorzugehen ist, stellen wir ebenfalls in der Folge vor.

– Best Practice: Internationale Studien

Die Planung und Durchführung von Studien in verschiedenen Gruppen beziehungsweise Kultur- und Sprachkreisen stellt grundsätzlich eine Herausforderung dar. Selbst die Anwendung der ›einfachen‹ Frage, die dem NPS-Ansatz zugrunde liegt, ist in diesem Kontext keineswegs trivial. Wir werden hier aus unserer mehrjährigen Praxis im Bereich internationaler NPS-Erhebungen von Problemen und best practice Ansätzen berichten.

2 Je nach eingeschätzter Wahrscheinlichkeit der Weiterempfehlung auf der 11-stufigen Skala werden die Befragten einer von drei Gruppen zugeordnet. Skalenergebnisse von 0 bis 6 werden als negative Antworten gewertet. Personen dieser Gruppe werden als Detraktoren bezeichnet – es ist damit zu rechnen, dass es durch diese Personen aufgrund weniger positiver Erfahrungen eher zu negativem Empfehlungsverhalten kommt. Personen mit den Skalenergebnissen 7 und 8 werden als neutrale Kunden bezeichnet. Die Bewertung ist zwar nicht negativ, es ist jedoch nicht davon auszugehen, dass eine aktive Weiterempfehlung erfolgt. Lediglich bei Personen die einen Skalenwert von 9 oder 10 wählen, wird davon ausgegangen, dass sie das bewertete Produkt oder Unternehmen aktiv weiterempfehlen. Diese Personengruppe wird als Promotoren bezeichnet. Laut Reichheld werden diese Personen zu Botschaftern für ein Unternehmen und reagieren gleichzeitig nur wenig auf Konkurrenzangebote. Der so genannte Net Promoter Score wird als Differenz des prozentualen Anteils der Promotoren und der Detraktoren ermittelt (Reichheld, 2003).

A. Vergleich mit Konkurrenten

Um mittels des NPS Vorhersagen über die mittel- und langfristige Marktentwicklung treffen zu können, gilt es, neben dem eigenen NPS Wert auch den der Konkurrenten zu betrachten. Es reicht also nicht aus, die eigenen Kunden hinsichtlich der Weiterempfehlungsbereitschaft zu befragen. Für eine aussagekräftige Bewertung der eigenen Position müssen auch die Kunden der (wichtigsten) Konkurrenten am Markt zu ihrer Weiterempfehlungsbereitschaft befragt werden. Hintergrund dieser Überlegung ist, dass die Markenempfehlung durch Kunden Einfluss darauf hat, wie sich potenzielle Kunden zukünftig verhalten. Diese potenziellen Kunden erhalten – ebenso wie eigene Kunden – auch Empfehlungen von Kunden der Konkurrenzmarken. Zufriedene Kunden fungieren »als Werbeträger für eine Marke indem sie von ihren positiven Erfahrungen [...] berichten. Diese Form der kostengünstigen Direktwerbung ist in Zeiten des »information overload«, in denen die Effizienz der klassischen Werbung abnimmt, von enormer Bedeutung. Kaufentscheidungen werden vor allem aufgrund von Empfehlungen vertrauter Personen getroffen« (Stahl, Binder, Cosler & Scharioth, 1998).

Die Notwendigkeit der zusätzlichen Konkurrenzanalyse bedeutet in der Regel einen erheblichen Mehraufwand bei der Durchführung einer (NPS-)Studie. Während über die eigenen Kunden (in der Regel) zahlreiche Informationen vorliegen und Kontaktwege vorhanden sind, fehlt dies in Bezug auf Konkurrenzkunden. Für ein Unternehmen ist es dann kaum mehr sinnvoll möglich, eine Befragung des gesamten Marktes selber durchzuführen. Hier empfiehlt sich der Rückgriff auf ein unabhängiges, neutrales Institut, welches Kunden anderer Unternehmen befragen kann – auch ohne Reaktanz auszulösen. Um eine einheitliche Erhebung sicherzustellen und NPS-Werte sinnvoll miteinander in Beziehung setzen zu können, sollten eigene Kunden und Konkurrenzkunden im Rahmen derselben Studie befragt werden. Selbstverständlich darf dann bei einer solchen Befragung der Auftraggeber – zumindest zu Beginn – nicht genannt werden.

Es stellt sich die Frage, wie man an eine repräsentative Stichprobe von Kunden aller relevanten Konkurrenten gelangen kann. Je kleiner die Marktdurchdringung der Produktgruppe ist, desto aufwendiger wird das Sampling. Bei einem Gut, welches in den meisten Haushalten (hier: B2C) vorhanden ist, z.B. ein PKW oder eine Kaffeemaschine, lässt sich eine repräsentative Haushaltsbefragung vergleichsweise einfach realisieren, wobei zu Beginn der Befragung nach dem Besitz des relevanten Produktes gescreent wird. Ist das Produkt im Haushalt vorhanden, lässt sich die Marke erfragen, für die schließlich die Weiterempfehlungsbereitschaft erfasst wird. Problematisch wird es dann, wenn das Produkt (z.B. ein PKW) mehrfach und gegebenenfalls von verschiedenen Marken im Haushalt vorhanden ist. Hier bietet sich die Möglichkeit an, die Weiterempfehlungsbereitschaft hinsichtlich aller

genutzten Marken zu erfragen oder sich auf die hauptsächlich genutzte Marke zu beschränken.

Der Vorteil des oben dargestellten Ansatzes einer repräsentativen Befragung, die sich so auch auf Unternehmensbefragungen übertragen lässt, besteht darin, dass man gleichzeitig auch Informationen über Marktanteile erfasst. Dabei lässt sich zwischen der Verteilung der Marken im Besitz der Kunden zum einen und der Verteilung der hauptsächlich genutzten Marken zum anderen unterscheiden. So können in einem Haushalt durchaus drei (oder mehr) Fernseher unterschiedlicher Marken vorhanden sein, hauptsächlich genutzt wird aber gegebenenfalls nur eines der Geräte. Diese Hauptnutzung kann sich damit von den verkauften Marktanteilen unterscheiden, ist aber unserer Meinung nach speziell im Bereich Kundenzufriedenheit und Weiterempfehlungsbereitschaft zentraler. In einigen Bereichen können sich die beiden Verteilungen erheblich voneinander unterscheiden. Im Bereich Mobilfunk beispielsweise zeigen die von den Providern berichteten Zahlen, dass nicht wenige Personen über mehr als eine Mobilfunkkarte (SIM) verfügen. Teilweise werden solche Karten recht offensiv in Kombination mit Mobilfunkgeräten auf den Markt gedrängt. Das heißt aber noch nicht, dass diese Karten auch regelmäßig genutzt werden. Hier unterscheiden sich die Marktanteile der Provider hinsichtlich der SIM-Karten im Markt deutlich von den »hauptsächlich« genutzten Marken. Es ist jedoch zu erwarten, dass eine Empfehlung an Freunde und Bekannte vor allem für die Marke ausgesprochen wird (beziehungsweise ausgesprochen werden kann), welche man selbst häufig nutzt.

Wer sich also umgekehrt überlegt, beispielsweise einen Mobilfunkvertrag abzuschließen oder seinen alten zu verlängern, wird sich eben auch bei Kunden anderer Provider informieren. Je höher der NPS-Wert der Konkurrenten ist, desto größer ist die Wahrscheinlichkeit, dass ein Unternehmen seine Kunden an den Konkurrenten verliert.

Durch die Kenntnis des eigenen NPS-Wertes ist also nichts gewonnen, solange man ihn nicht mit denen der Konkurrenten vergleicht. Ein NPS-Wert von z.B. +50% kann als Indikator für eine positive Marktentwicklung gesehen werden, wenn alle Konkurrenten niedrigere Werte haben. Haben die Konkurrenten jedoch höhere Werte, so ist zu befürchten, dass diese schneller wachsen oder sogar Kunden zum Wechsel bewegen. Umgekehrt kann also auch ein niedriger oder gar leicht negativer NPS-Wert auf ein Wachstum hindeuten, wenn das Marktumfeld schlechtere NPS-Werte aufweist.

Zum Vergleich des eigenen NPS-Wertes empfiehlt sich nicht nur der Vergleich mit jedem einzelnen Konkurrenten, sondern vor allem der mit dem mittleren NPS des Marktes. Der »Empfehlungsdruck«, der auf eine einzelne Person wirkt, hängt auch von den Marktanteilen der verschiedenen Marken ab. Wenn eine Marke nur einen Marktanteil von 5% hat, so wird ein potenzieller Kunde entsprechend selten mit einer Empfehlung für dieses Produkt

(bzw. einer Warnung vor diesem Produkt) in Kontakt kommen. Bei der Bestimmung des mittleren NPS-Wertes der Konkurrenten gilt es, dies zu berücksichtigen. Daher sollte von einer einfachen Berechnung des arithmetischen Mittels der NPS-Werte abgesehen werden. Vielmehr sollten diese gewichtet mit ihrem relativen Marktanteil berücksichtigt werden.

Für die Bestimmung des Gesamtmittelwertes des NPS wird der NPS jeder Marke mit seinem Marktanteil multipliziert und darüber dann die Summe gebildet. Es finden sich z.B. für drei Anbieter A, B und C Marktanteile von $m_A = .55$, $m_B = .30$ und $m_C = .15$ und NPS-Werte von $NPS_A = +20\%$, $NPS_B = +15\%$ und $NPS_C = +40\%$. Dann ist der Gesamtmittelwert NPS_{Gesamt} des NPS über alle Marken:

$$NPS_{Gesamt} = m_A * NPS_A + m_B * NPS_B + m_C * NPS_C = .55 * 20\% + .30 * 15\% + .15 * 40\% = 21,5\%$$

Marke A liegt mit ihrem NPS also leicht unter dem Durchschnitt, B deutlich unter und Marke C deutlich über diesem.

Für die genauere Marktprognose der Marke A wird der Mittelwert $NPS_{Kon(A)}$ nur der Konkurrenten von A berechnet.

$$NPS_{Kon(A)} = (m_B * NPS_B + m_C * NPS_C) / (m_B + m_C) = (.30 * 15\% + .15 * 40\%) / (.30 + .15) = 23,3\%$$

Der mittlere NPS-Abstand Delta von Marke A zu ihren Konkurrenten beträgt also $-3,3\%$:

$$Delta_A = NPS_A - NPS_{Kon(A)} = +20\% - 23,3\% = -3,3\%$$

Ziel zukünftiger (Marketing-)Maßnahmen ist es, diese Differenz $Delta_A$ zu minimieren oder die Konkurrenten sogar zu überholen.

Grundsätzlich stellt sich die Frage, ob die gefundene Differenz tatsächlich statistisch bedeutsam ist. Die gemessenen NPS-Werte beruhen (typischerweise) auf einer Stichprobe und sind damit immer fehlerbehaftet. Speziell bei wiederholten Messungen, denen sich der nächste Abschnitt widmet, gilt es zu entscheiden, ob zwischen zwei Messzeitpunkten gefundene Veränderungen zufällig sind oder sich dahinter eine wirkliche Verbesserung (oder Verschlechterung) der eigenen Marke verbirgt.

B. Vergleich über die Zeit

Für eine kennzifferngeleitete Unternehmenssteuerung mittels NPS ist es erforderlich den eigenen NPS-Wert und den der Konkurrenten regelmäßig zu erheben, so dass sich eine Reihe von Messungen zu verschiedenen Zeitpunkten i ($NPS_{A[i]}$, $NPS_{Kon(A)[i]}$, $Delta_{A[i]}$ usw.) ergeben. Nur so lassen sich frühzeitig Chancen und Risiken erkennen. Dabei sollte, wie oben ausgeführt, nicht der eigene absolute NPS-Wert, sondern vielmehr der mittlere Abstand (Delta) zu den Konkurrenten verfolgt werden. Die Verbesserung dieses Deltas zwischen zwei Messzeitpunkten i und j kann als direkt und zeitnah messbares Marketingziel definiert werden.

$$\text{diffDelta}_{A[i;j]} = \text{Delta}_{A[j]} - \text{Delta}_{A[i]}$$

Als eine solche Schlüsselkennzahl (KPI, key performance indicator) muss diese Veränderung $\text{diffDelta}_{A[i;j]}$ statistisch abgesichert werden: Die beobachteten Veränderungen sollten gegen den Zufall geprüft werden. Zwar ist die Freude groß, wenn sich das Delta_A von Messzeitpunkt i nach $i+1$ beispielsweise um 2% verbessert. Die Frage nach der statistischen Bedeutsamkeit wird dann gerne vernachlässigt. Wenn sich das Delta_A jedoch zum nächsten Zeitpunkt $i+2$ wieder um 3% verschlechtert, wird leicht nachvollziehbar, dass es sinnvoll ist, stets die statistische Signifikanz der Veränderung zu betrachten.

Die Verteilung von Delta_A ist beeinflusst durch a) die Größe der Stichprobe N , b) die verschiedenen Marktanteile und c) die Verteilung der Promotoren und Distraktoren jeder Marke. Die analytische Bestimmung der Verteilung ist nicht trivial, daher verwenden wir in unserer Forschungspraxis hierfür einen Monte-Carlo-Simulationsansatz. Dabei nutzen wir die bei einer ersten Basismessung bestimmten Werte als Parameter und simulieren dann 10.000 Stichproben derselben Größe. Für jede dieser Stichproben bestimmen wir unsere relevante Statistik Delta_A und können so beliebige Konfidenzintervalle bestimmen.

Typischerweise arbeiten wir mit dem 95%-Konfidenzintervall. Dieses Konfidenzintervalls $\text{KI}_{95\%(\text{Delta}_A)}$ gibt den Bereich an, in dem sich in 95% der Fälle das berechnete Delta_A befindet. Für unser Beispiel von oben mit einem gefundenen Delta_A von -3,3% könnte dieser Bereich zwischen -5,3% und -1,3% liegen. So lange bei den nächsten Messungen Delta_A in diesem Bereich bleibt, gehen wir davon aus, dass keine bedeutsame Veränderung stattgefunden hat, sondern lediglich Zufallsschwankungen vorliegen. Erst bei größeren Veränderungen gehen wir von einer bedeutsamen Verbesserung oder Verschlechterung des NPS-Wertes der Marke aus. Die Breite $\text{BrKI}_{95\%(\text{Delta}_A)}$ des Konfidenzintervalls $\text{KI}_{95\%(\text{Delta}_A)}$ beträgt also 4%.

Als Maß für die Zielerreichung betrachten wir also nicht den absoluten Wert $\text{diffDelta}_{A[i;j]}$, der Änderung von Delta_A zwischen zwei Zeitpunkten i und j , sondern relativieren diese an der Breite $\text{BrKI}_{95\%(\text{Delta}_A)}$ des Konfidenzintervalls³:

$$\text{Kdif}_{A[i;j]} = \text{diffDelta}_{A[i;j]} / \text{BrKI}_{95\%(\text{Delta}_A)}$$

Erst wenn der Quotient $\text{Kdif}_{A[i;j]}$ absolut größer 1 ist, gehen wir von einer signifikanten Veränderung aus. Ein weiterer Vorteil dieses Quotienten ist es, dass damit auch Entwicklungen in verschiedenen Bereichen sinnvoll vergli-

3 Statistisch korrekt müsste das Konfidenzintervall für jeden Messzeitpunkt i neu simuliert werden. Dies würde jedoch die unternehmensinterne Kommunikation erschweren, da mit jeder Messung neue Zielvorgaben erfolgten. Wir errechnen daher nur zum 1. Zeitpunkt einen Basiswert und bestimmen damit einmal das $\text{KI}_{95\%(\text{Delta}_A)}$ für die folgenden Messungen.

chen werden können. Somit kann über verschiedene Bereiche ein einheitliches Zielerreichungsprogramm implementiert werden, wie im nächsten Abschnitt ausgeführt.

C. Vergleich der Entwicklung in verschiedenen Bereichen

Noch komplexer wird die Situation, wenn die Entwicklung des NPS eines Unternehmens in verschiedenen Bereichen verglichen werden soll. Denkbar ist beispielsweise, dass eine einheitliche Zielerreichungslogik für verschiedene regionale Einheiten (Bundesländer, Regionen oder Länder) oder für unterschiedliche Zielgruppen oder Produkte entwickelt werden soll⁴. Die Ausgangssituation wird sich hier zwischen den verschiedenen Bereichen deutlich unterscheiden, ein direkter Vergleich der NPS Werte verbietet sich daher. Auch hier muss wieder auf die Entwicklung der Deltas in den verschiedenen Bereichen zurückgegriffen werden.

Doch selbst die Veränderungen der Deltas können nicht eins zu eins miteinander verglichen werden, sondern müssen standardisiert werden. Hier bietet sich die Relativierung an der Breite des jeweiligen Konfidenzintervalls, wie oben beschrieben, an.

Für dieses Vorgehen werden zunächst in allen Bereichen Basismessungen durchgeführt und die mittleren Abstände Delta zu den Konkurrenten in den jeweiligen Bereichen (oder Märkten) berechnet. Zu überprüfendes Ziel ist es, diese Abstände (weiter) zu verbessern. Um die verschiedenen Bereiche vergleichen zu können, werden für jeden Bereich (oder Markt) zunächst die Konfidenzintervalle für Delta simuliert. Als marktübergreifende Kennziffer für die Änderungen über die Zeit wird die Änderung von Delta relativiert an der Breite des entsprechenden Konfidenzintervalls definiert. Veränderungen über die Zeit können dann über alle Bereiche (Märkte) in der gleichen Metrik $Kdif_{A[i..j]}$ betrachtet werden. Als Ziel kann vereinbart werden, dass sich die Differenz zu den Konkurrenten z.B. um die Breite eines Konfidenzintervalls oder mehr verbessern soll oder z.B. nicht um mehr als die Breite eines Konfidenzintervalls verschlechtern darf. Die Zielerreichung kann dann auch sinnvoll abgestuft werden und Boni oder Mali vereinbart werden.

D. Best Practice: Internationale Studien

Zunehmende Internationalisierung wirtschaftlicher Tätigkeit erfordert auch internationales Denken hinsichtlich der Vorbereitung von Marketingentscheidungen (Berndt, Fantapié Altobelli & Sander, 1999). Unternehmertätigkeit im internationalen Umfeld setzt somit internationale (Markt-)For-

4 Zur Problematik der Studienplanung und Datenerhebung siehe Abschnitt D im vorliegenden Artikel.

schung unverzichtbar voraus (u.a. Johnson, 1998; Holzmüller & Woitschläger, 2007).

Multinationale Forschungstätigkeit weist eine lange Tradition auf, in welcher einerseits das Problembewusstsein geschärft wurde und andererseits eine stetige methodische Weiterentwicklung stattfand. Dennoch weist internationale Forschungstätigkeit häufig methodische Schwächen auf oder widmet der internationalen Planung beziehungsweise Durchführung von Studien zu wenig Aufmerksamkeit. Es wurde bereits darauf hingewiesen, dass ein direkter Vergleich von Kennwerten verschiedener Gruppen, Regionen oder Länder aufgrund unterschiedlicher Ausgangssituationen nicht ohne weiteres möglich ist. Ein unreflektierter Vergleich birgt die Gefahr, *irrelevante* Gruppenunterschiede zu identifizieren, die allein Resultat mangelnder Vergleichbarkeit von Ausgangssituation und Erhebungsinstrument und damit mangelnder Vorsicht bei der Instrumententwicklung und Studienplanung sind. Diesen potenziellen Fehlerquellen gilt es bereits bei der Studienkonzeption, also lange vor der Datenerhebung sowie Analyse und Interpretation von Ergebnissen, zu begegnen.

Häufig werden internationale Forschungsprojekte »von einem Willenszentrum in einer Ausgangskultur angestoßen« (Ertl, 2001). So besteht die Gefahr, dass durch einen eingeschränkten Blickwinkel keine erschöpfende Erfassung des Forschungsgegenstandes erfolgt und so ein – für bestimmte partizipierende Untersuchungseinheiten – unvollständiges Bild entsteht. Interkulturelle Forschung bedarf daher einer grundsätzlichen Auseinandersetzung mit dem Umgang der Erfassung von Konstrukten in verschiedenen sozialen oder kulturellen Entitäten. Sie verfolgt in der Regel das Ziel, Konstrukte analog in verschiedenen Gruppen, Regionen oder Ländern zu erfassen und gegebenenfalls zu vergleichen. Voraussetzung für die analoge Erfassung von Konstrukten erfordert jedoch zunächst deren äquivalente Existenz⁵. Dass das Verständnis (aus Perspektive von nur einer Gruppe) trivial erscheinender Konstrukte in verschiedenen Gruppen voneinander abweicht, mag durch kulturelle Verzerrung des Verständnisses von Forschern rasch vernachlässigt werden (u.a. Hofstede, 1996; Braun & Harkness, 2005). Um der Herausforderung der Überwindung des Ethnozentrismus gerecht zu werden, nationale

5 Im Rahmen der Entwicklung internationaler Forschungsfragestellungen ist aus diesem Grund die Emic-Etic Diskussion zu führen. Die Begrifflichkeiten gehen auf den Linguisten Pike (1967) zurück, der diese von phonemic und phonetic ableitete. Etic-Orientierung beschreibt Forschung, die sich auf kulturübergreifende, universale Konstrukte bezieht, die multinational analog zugänglich sind. Demgegenüber wird bei der Emic-Orientierung eine einzigartige Existenz von Konstrukten innerhalb eines Kultursystems postuliert. Etic- und Emic-Orientierung lassen sich jedoch nicht als exklusive Extreme verstehen. Eine Kombination der Perspektiven zielt auf ein Maximum an Vergleichbarkeit unter Berücksichtigung nationaler Besonderheiten ab (Douglas & Craig, 2006).

Besonderheiten zu erkennen und zu berücksichtigen und um ein Befragungsinstrument zu entwickeln, welches sich zur Operationalisierung der Fragestellung im multinationalen Raum eignet, empfiehlt es sich daher, die Vorteile der Arbeit mit multinationalen Teams zu nutzen (Eastery-Smith & Malina, 1999). Auf diese Weise lässt sich ein Befragungsinstrument entwickeln, das der Diversität partizipierender Gruppen Rechnung trägt. Bei der Studienplanung ist darüber hinaus auf Standardisierung der Untersuchungsmethode sowie auf Vergleichbarkeit der zu untersuchenden Personengruppen einerseits und der Untersuchungssituation andererseits zu achten.

In der Praxis der NPS-Forschung wird in der Regel von einer pankulturellen Existenz des Konstrukts der Weiterempfehlungsbereitschaft ausgegangen. Das Konzept der NPS-Erfassung, welches ja die Generierung einer aussagekräftigen Kennzahl aufgrund einer singulären Frage unterstellt, birgt vergleichsweise wenig Potenzial für intersoziale oder interkulturelle Missverständnisse. Dennoch können wir aufgrund unserer Forschungserfahrung nur betonen, dass selbst die Übertragung der »one number you need to grow« (Reichheld, 2003) einen nicht zu vernachlässigenden Diskurs auslösen kann, der – dem in der Marktforschung nicht unüblichem Zeitdruck zum Trotz – geführt werden muss, um nutzbare Ergebnisse zu generieren. So ist auch die Übertragung dieser vermeintlich einfachen Frage von der Ausgangssprache in andere Sprachen nicht immer eindeutig. Wir konnten beispielsweise in unserer Forschungspraxis erleben, dass drei verschiedene Übersetzungen des Wortes »Weiterempfehlung« ins Ungarische angeboten wurden, die alle korrekt waren, sich jedoch im Wortlaut erheblich unterschieden. Häufig sind verschiedene Übersetzungen denkbar und es gilt sich frühzeitig mit sprachlichen Feinheiten auseinander zu setzen, um einerseits inhaltliche Übereinstimmung als Grundlage für Vergleichbarkeit sicher zu stellen und andererseits Verbindlichkeit für die verschiedenen in einer Studie zu bewertenden Länder zu garantieren. Gerade in internationalen Studien bietet sich andernfalls die Übersetzung als »Ausrede« für schlechte Ergebnisse an. Dies kann nicht nur zu erheblichen (vermeidbaren) Diskussionen führen, sondern mitunter dazu, dass Studienergebnisse angezweifelt werden oder Kennwerte nicht anerkannt werden.

E. Zusammenfassung

Im vorliegenden Artikel wird deutlich, dass die Erfassung von Unternehmenskennzahlen sowie deren Anwendung in der Praxis zahlreiche Fallstricke birgt. Selbst im Falle theoretisch wohldefinierter Auswertungslogik und einfacher Fragestellung, kann sich die Implementierung von Kennzahlen in die Praxis als komplex erweisen. Am Beispiel der Erfassung und Auswertung des Net Promoter Scores haben wir Nutzen und Probleme des Umgangs mit

Kennziffern zur Unternehmenssteuerung – im Marktumfeld, im Zeitvergleich und im internationalen Kontext – aufgezeigt und Lösungsansätze vorgestellt.

So ist generell vor einer unreflektierten beziehungsweise unrelativierten Nutzung von Schlüsselindikatoren wie dem NPS zu warnen. (Markt-)Forschungstätigkeit sowie die Interpretation von (Markt-)Forschungsergebnissen sind stets in den Rahmen der spezifischen Marktsituation einzubetten. Zu warnen ist nicht nur vor dem unachtsamen Vergleich von Forschungsergebnissen verschiedener Untersuchungsgruppen, sondern auch davor, Instrumente, die zur Erfassung eines Konstrukts entwickelt wurden, sorglos zur Erfassung des (vermeintlich) selben Konstrukts auf andere Gruppen oder Kulturen zu übertragen.

Literatur

- BERNDT, R., Fantapié Altobelli, C., & Sander, M. (1999). *Internationales Marketing-Management*. Berlin: Springer.
- BRAUN, M. & Harkness, J. A. (2005). Text and context: Challenges to comparability in survey questions. In J. A. Harkness (Ed.), *Methodological aspects in cross-national research* (Vol. 11, p. 95-107). Mannheim: ZUMA.
- DOUGLAS, S. P. & Craig, C. S. (2006). On improving the conceptual foundations of international marketing research. *Journal of International Marketing*, 14 (1), 1-22.
- EASTERBY-Smith, M. & Malina, D. (1999). Cross-Cultural collaborative research: Toward reflexivity. *Academy of Management Journal*, 42 (1), 76-86.
- ERTL, R. (2001). *Nationale Kundenbarometer im Vergleich. Theorie – Empirie – Gestaltungsvorschläge*. München: Gabler.
- HOFSTEDE, G. (1996). *Culture and organizations: Software of the mind; Intercultural Cooperation and its importance for survival*. New York: McGraw-Hill.
- HOLZMÜLLER, H. H. & Woisetschlager, D. M. (2007). Herausforderungen in internationalen Forschungsprojekten. In A. Herrmann, C. Homburg & M. Klarmann (Hrsg.), *Handbuch Marktforschung: Methoden – Anwendungen – Praxisbeispiele* (3. Auflage, S. 123-150). Wiesbaden: Gabler.
- JOHNSON, T. P. (1998). Approaches to equivalence in cross-cultural and cross-national research. In J. A. Harkness (Ed.), *Cross-Cultural Survey Equivalence* (Vol. 3, p. 1-40). Mannheim: ZUMA.
- PIKE, K. L. (1967). *Language in Relation to a Unified Theory of the Structure of Human Behavior* (2., revised edition). The Hague: Mouton.
- REICHELIED, F. F. (2003). The one number you need to grow. *Harvard Business Review*, 81 (12), 46-54.
- STAHL, M., Binder, G., Cosler, D., & Scharioth, J. (1998). *TRI:M-Studie zur Kundenzufriedenheit (Mehrfachkunden) 1997*. Bonn: Informationszentrum Sozialwissenschaften

Michael Paulus

Using NPS to play offense

The General Electric Company:

The General Electric Company is one of the leading companies in the world. More than 300.000 employees are generating a revenue of over \$160bn in the business segments Healthcare, Infrastructure, Industrial, NBC Universal, Commercial and Consumer Finance. GE is a truly multinational enterprise employing more than 70.000 employees in Europe.

GE Healthcare generates a revenue of \$20bn in 2007 and has 40.000 employees. Products include Diagnostic Imaging Products like CT, MR and X-ray as well as Ultrasound equipment. GE Healthcare also covers IT-systems, Biosciences and Intensive Care products.

The article covers GE's experience in implementing and running NPS globally as well as in the German and EMEA-market for the Clinical Systems Devices Service business. Service includes preventive maintenance, check-ups, remote monitoring and repair of the medical equipment. Customers are mainly hospitals but also private doctors.

NPS & the corporate environment

GE was one of the first companies globally to implement and use the Net Promoter Philosophy developed by Fred Reichheld and Satmetrix. The first business unit in GE Healthcare started as early as 2004 to implement the Net Promoter Score for their Service business to look at the Transactional NPS. First customer measurements were done and the program was widely communicated and key people were trained on the philosophy.

Jeff Immelt, CEO of General Electric, saw in 2005 the first promising results from GE Healthcare's experience with NPS from 2004. He saw that improving the NPS led to an increase in profitable growth and that NPS was allowing to implement a customer centric organization. At that time he recommended the other five GE businesses to look at the successes from GE Healthcare's implementation of NPS. In 2005 GE Healthcare was globally rolling out NPS and was showing more astonishing results directly linked to the NPS implementation. 2005 was also used to standardize the processes around NPS and enable a global NPS survey rigor. In 2005 the active Management of the »Voice of the Customer (VOC)« started. In early 2006 Jeff Immelt completely bought into the NPS-philosophy and asked all GE businesses to implement NPS as a key metric. By that time the Leadership was

showing ownership and engagement. Systems to enable a VOC transparency enabled driving »Big bets« as a recipe to create promoters. Now all GE businesses implemented the NPS and the underlying Operating System. 2007 is showing the results of the »Big Bets« implementation and the financial impact of NPS can be demonstrated. GE is now reaching the next level of understanding customer needs and will be able to further improve the NPS.

In former times GE was well known for Growth through acquisitions. Jeff Immelt's strategy is more focused around organic growth rather than acquisitions and therefore a fundamental shift in strategy asking for new tools and mechanisms. GE's goal is to maintain a growth rate twice as high as the global growth in the GDP. One key element to enable those significant organic growth rates was the implementation of the NPS framework that is allowing to realize this goal.

NPS Operating Mechanism & Framework

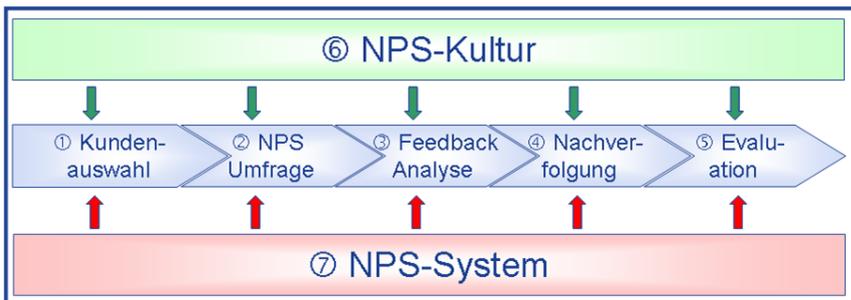
NPS is a simple metric. It is easy to understand and easy to see progress or drawbacks in customer enthusiasm. This is what makes NPS unique and powerful. It can be included in the Operating Mechanisms for all company levels from Field-Force to Top-Management. It always brings the Customer Experience in the center of the discussion and allows prioritization of customer needs, benchmarking and best-practice sharing.

Nevertheless NPS requires a lot of work to keep it running – and make it a success. Centered around the NPS metric the organization requires a complete Operating System and framework to make it work. It starts with a continuous education and training of the workforce to have the same common understanding of NPS. It also requires a customer segmentation to be able to cluster customer responses. Systems enabling a smooth process from service event to NPS-survey up until the possibility to manage customer responses and escalations are required. With data from those systems in hand a thorough – and time-consuming analysis leads to strategic priorities for field-force as well as management. Implementing an NPS survey rigor is important to make sure that all entities measure the same way to make results comparable and to allow a constructive benchmarking. Also, very important is a continuous Leadership ownership to make NPS one of the organizations key priorities. Tying the NPS progress to Goals & Objectives is an implicitness for Management but difficult to implement for the Field force, as usually there is an involvement of unions and workers council. Finally an evaluation is required to proof that growth in customer enthusiasm is linked to profitable growth continuously proofing the NPS promise.

NPS & the customer experience

Basically all companies put their customers first. This is what can be found in company brochures, mission statements, customer presentations and so on. However, the larger the organization grows the more difficult it is to truly live this commitment. The tendency towards increased bureaucracy to keep the organization running as well as global internal processes to enable efficient economies of scale are leaving the customer experience behind. Also, in a finance driven world the customer can not be measured and quantified in a dashboard leaving him behind in the organizational setup. With the help of NPS the »Customer first« experience can be made tangible and can be included in the Management Operating Mechanisms.

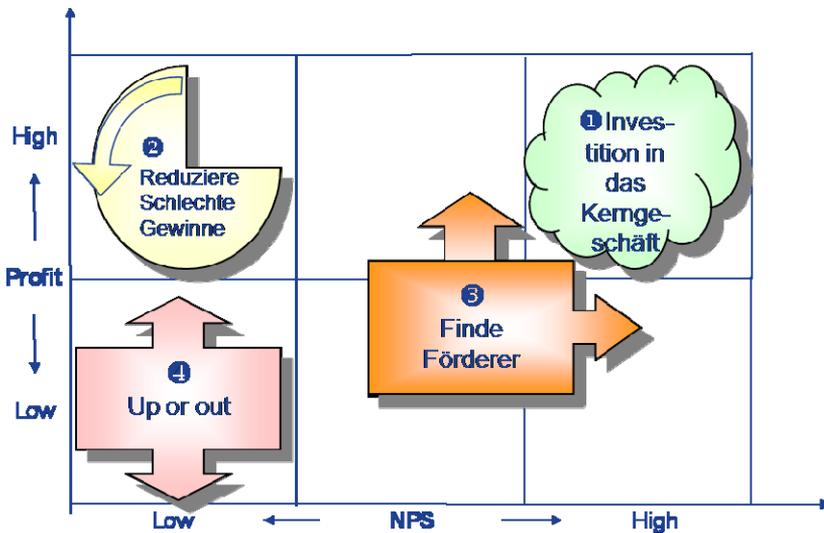
The GE Healthcare Clinical Systems Service business did a high 5-digit number of customer surveys in 2006 and will do the same amount again in 2007. What were the specific lessons learned from those responses and what did GE do to respond to customer's needs. Most surprising was the fact that responses were pretty consistent from different markets. Customers in North America had the same priorities as customers in Eastern Europe, Asia or Western Europe. Also very interesting was the 2nd biggest area for improvement – communication. Customers mention in their NPS-response to a tactical field service event that GE is doing a very good technical job when performing the service. There were no or little improvements to be seen by the customer. The area of improvement that was mentioned was that the technical staff did not communicate well enough with the customer about the service event. This included the scheduling of the event, the actual work at the customer site as well as the debrief of the service event with the customer. One of the main actions implemented was that now continuously soft-skills training and coaching is part of the training programs to improve the communication. With the continuous measurement of NPS-responses it was possible to see a decline in customers-responses on communication issues shortly after the soft-skills training was implemented.



When GE first implemented NPS numerous points were addressed by customers that had an impact on the customer satisfaction. Therefore GE initi-

ated a high number of projects using the internal Six-Sigma organization to tackle all those areas of improvement. Apart from resource constraints it turned out that customers were not feeling many of the tiny improvements. Therefore the strategy was changed and GE is now focusing on a few game-changing projects that will have a significant impact which will also be felt by the customer.

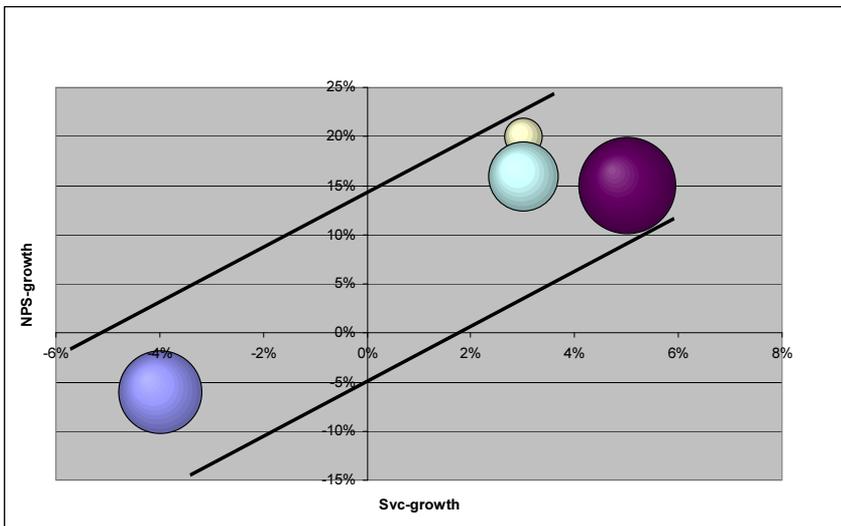
A thorough customer segmentation is key for the correct interpretation of the customer responses. Without segmentation contradicting requirements from customers can not be put into an improvement-plan. One example is that one customer category requires speedy repairs whereas another customer segment has no interest in fast repairs as enough backup-machines are available. With a segmentation approach each customer segment can be profiled correctly and strategies for those customer segments can be built. The segmentation approach also allows you to focus on growing customers with high profit and high NPS and 2nd work on reducing customers with a high profit but low NPS.



One key-learning was to better understand the important role that employees play in putting the customer ahead of the organization. A strong correlation was found between customer and employee satisfaction. Basically an employee will only be able to make a customer happy if the employee himself is happy as an individual within the GE-organization. GE found many different levels of employee satisfaction in different management teams and organizations and therefore GE introduced the employee-NPS within the organization to find out what needs to be done in the individual units to improve the employee loyalty and enthusiasm. According to the NPS-

philosophy the question to be asked on a regular basis is »...would you recommend GE as an employer«. The analysis of the results is now used to continually improve the score.

The basic concept with NPS is that an increase in customer loyalty leads to increased profitable growth. GE was and is continuously verifying that the concept works within the businesses reviewed. The experience is that not all organizations within GE develop the same, some grow and some shrink, however there is a clear correlation between NPS and growth. The growing businesses show a positive improvement in NPS and the shrinking organizations show a negative trend in NPS. This is supported by external surveys correlating with the NPS-trend and with the development of the market-share. One sample outcome (graph) was that a strong correlation was found between a shrinking business and a low (negative) NPS and with other growing business and a high positive and improving NPS. The graph shows four country organizations in comparison. The bubble size represents the size of the business.



This evaluation of the NPS framework is crucial and critical to the success of the program. Management, driven by financial metrics, needs to see and understand the positive link between NPS and profitable growth to support the program.

Summary

GE implemented NPS as early as 2004 and has a vast experience in introducing, running and finetuning NPS programs in all business units globally.

NPS allows GE to grow organically faster than the GDP. Apart from the NPS-metric a sophisticated framework was established to support processes and Operating Mechanisms. With those tools the Customer experience was included in Management Reviews and a follow up on Customer's strategic priorities was possible. Closed loop processes allow for a detractor follow-up and a continuous program evaluation is done including best-practice sharing and benchmarking on a global scale.

To be successful with the NPS-implementation the Employee buy-in from Top-Management to Field-Force is required. Designing a few »big bets« changes based on customer feedback to develop the organization towards a customer-centric organization must be based on a profound customer segmentation and is essential for a winning strategy.

Astrid Mayerböck

Bedarfsanalyse eines Gründerzentrums an der Universität Bonn

Befragung zum Gründungsgeschehen und zur Gründungserfahrung bei Mitarbeitern der Universität Bonn

Hintergrund der durchgeführten Untersuchung ist das Interesse der Stadt Bonn, ein Gründerzentrum zu errichten, das insbesondere Raum für Ausgründungen aus der Universität Bonn heraus bieten soll. Im Rahmen der Zielvereinbarung zur Studienreform zwischen dem Land Nordrhein-Westfalen und der Universität Bonn sollen Qualität von Aus- und Weiterbildung sowie Wettbewerbsfähigkeit der Universität gefördert werden. Zur Wettbewerbsfähigkeit tragen unter anderem Kooperationen und Vernetzungen zwischen Wirtschaft und Universität bei. In diesem Zusammenhang gewinnen gerade Ausgründungen aus der Universität heraus an Bedeutung. Die Universität Bonn ist bemüht, ihre Mitarbeiterinnen und Mitarbeiter im Hinblick auf potenzielle Ausgründungen zu unterstützen und zu fördern. Auf Anregung der Transferstelle wurden daher in einer Bedarfsanalyse Mitarbeiterinnen und Mitarbeiter bezüglich ihres Gründungspotenzials und -interesses befragt. Zudem wurden Gründerinnen und Gründer so genannter Spinoff-Unternehmen – Ausgründungen aus der Universität – zu ihren Erfahrungen im Rahmen der Unternehmensgründung befragt. Um zu einer übergreifenden Bewertung zu kommen, wurden die Dekane aller Fakultäten zu einem persönlichen Gespräch eingeladen. Sie wurden gebeten, Aussagen zum Gründungsgeschehen in ihrem Fachbereich bzw. zu besonderen Anforderungen aus der jeweiligen Fakultät heraus zu treffen. Auf diese Weise konnten die Ergebnisse der Befragung abgerundet bzw. verifiziert werden.

Ziel der Befragung war, das Gründungsinteresse und den Bedarf der Universitätsmitarbeiterinnen und -mitarbeiter zu erfassen, um dies gegebenenfalls in einem Gründerzentrum der Stadt Bonn abbilden zu können.

Feldzeit und Stichprobe

Die Datenerhebung erfolgte in Form einer Online-Befragung in der Zeit vom 12.07.2006 bis zum 17.08.2006. Über das Rechenzentrum wurden insgesamt 5.575 E-Mail Einladungen zur Teilnahme an der Befragung versandt. Wie viele dieser Accounts jedoch aktiv genutzt wurden, lässt sich nicht prüfen. Zudem wurden 49 Gründerinnen und Gründer so genannter Spinoff-

Unternehmen angeschrieben, die aus der Universität Bonn heraus entstanden. Insgesamt beteiligten sich $n = 339$ Mitarbeiterinnen und Mitarbeiter der Universität sowie $n = 15$ Gründerinnen und Gründer von Spinoff-Unternehmen und nahmen Stellung zu folgenden Themen:

- Gründungsabsichten bzw. Gründungserfahrungen
- Kenntnisse von Beratungs- und Förderprogrammen der Universität und der Stadt Bonn hinsichtlich Unternehmensgründungen
- Interesse an einem Gründerzentrum sowie sinnvolle Ausstattung.

Darüber hinaus erfolgte die Befragung der Dekane zum Gründungsgeschehen bzw. Gründungsinteresse innerhalb ihrer Fakultät im persönlichen Gespräch.

Gründungspotenzial

Das Gründungsgeschehen bzw. -interesse der Mitarbeiterinnen und Mitarbeiter ist überraschend hoch. In der Vergangenheit waren knapp 11% der Befragten bereits an mindestens einer Unternehmensgründung beteiligt (das entspricht absolut 36 Personen). Fast 6% (19 Personen) befassen sich aktuell mit einer Gründung und 26% (84 Personen) haben sich schon einmal ernsthaft überlegt, in Zukunft ein Unternehmen zu gründen. Es sind vor allem Mitarbeiterinnen und Mitarbeiter in Forschung und Lehre (im Ggs. zu nicht-wissenschaftlichen Mitarbeiterinnen und Mitarbeitern), die bereits Gründungserfahrungen sammeln konnten bzw. aktuell gründen oder sich dies für die Zukunft vorstellen können. Die Gründungsidee entwickelten 51% der Mitarbeiter mit Gründungserfahrungen im Rahmen ihrer Tätigkeit bzw. im Rahmen ihres Studiums an einer Hochschule. Auch die Idee für aktuell laufende bzw. zukünftig geplante Gründungen entstand bei 60% der Befragten im Zusammenhang mit Tätigkeit oder Studium an der Universität.

Betrachtet man das Gründungspotenzial der verschiedenen Fakultäten, wird deutlich, dass das Thema Unternehmensgründung vor allem in der Mathematisch-Naturwissenschaftlichen Fakultät eine größere Rolle spielt. Sowohl hinsichtlich bereits durchgeführter Gründungen als auch bezüglich aktuellen Gründungsgeschehens und Gründungsvorhabens ist diese Gruppe zahlenmäßig am stärksten. Im Gegensatz dazu spielt weder in der Katholisch- noch in der Evangelisch-Theologischen Fakultät das Thema Unternehmensgründung eine Rolle. Die übrigen Fakultäten stehen der Thematik durchaus aufgeschlossen gegenüber, wenn auch das absolute Gründungsgeschehen deutlich niedriger ist als in der Mathematisch-Naturwissenschaftlichen Fakultät. Auch die unterschiedliche Beteiligung aus den verschiedenen Fakultäten, die sicherlich nicht zuletzt mit Interesse an der Thematik zusammenhängt sowie die Aussage der befragten Dekane spiegelt diese Verteilung des Gründungspotenzials aus den verschiedenen Fachbereichen wider.

Gründungsinhalt und Ablauf

Von den Gründern benannte Gründungsgegenstände wurden vor dem Hintergrund ihrer Eignung für die Ansiedlung in einem Gründerzentrum betrachtet. Bei bereits gegründeten Unternehmen handelt es sich im Wesentlichen um Gründungen im Dienstleistungsbereich, die generell als gut geeignet für die Ansiedlung in einem Gründerzentrum gelten. Die genannten Dienstleistungsgegenstände wurden wiederum nach ihrer Eignung hinsichtlich der Nutzung eines Gründerzentrums (siehe Abbildung 1) differenziert. »Hochwertige Gründungen« sind hierbei solche Gründungen, bei denen eine Ansiedlung in einem Gründerzentrum problemlos möglich ist. Hierbei handelt es sich vor allem um Beratungsangebote in unterschiedlichen Fachbereichen, aber auch speziellere Gründungsformen, beispielsweise im Bereich der Geoinformatik etc. Die hier als »sonstige Dienstleistungen« gekennzeichneten Gründungen sind demgegenüber für eine Ansiedlung in einem Gründerzentrum weniger gut geeignet, wengleich auch hier wertschöpfende Aktivitäten vorliegen (beispielsweise Lebensberatung, aber auch Apotheken, Arztpraxen, Elektrofachgeschäfte etc.).

Aktuell erfolgende Gründungen beziehen sich neben dem Dienstleistungssektor sehr häufig auf den Bereich Software und IT.

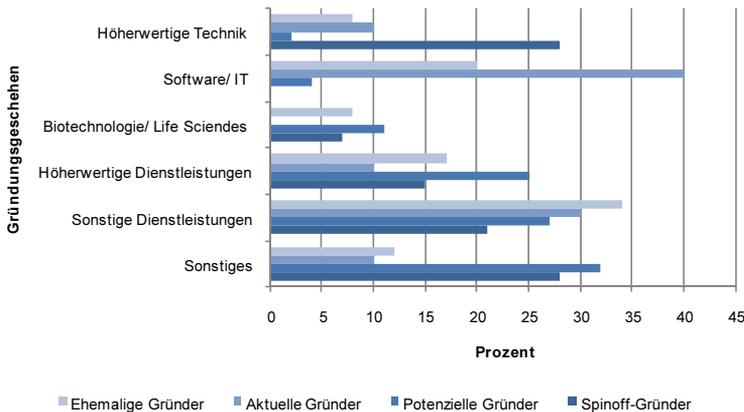


Abbildung 1: Gründungsgegenstand

Die Marktetablierung bereits gegründeter Unternehmen (durch Mitarbeiter der Universität) nahm durchschnittlich nur 2,7 Jahre in Anspruch – 41% der Unternehmen etablierten sich sogar innerhalb des ersten Geschäftsjahres. Die befragten Spinoffs benötigten durchschnittlich 2 Jahre zur Marktetablierung.

Für Unternehmen, die aktuell durch Universitätsmitarbeiter gegründet werden bzw. für die Zukunft geplant sind, erwarten die Gründer eine etwas längere Etablierungsphase. Nur 22% der potenziellen Gründerinnen und

Gründer erwarten, dass diese Phase innerhalb eines Jahres abgeschlossen sein wird.

Nutzung universitärer und außeruniversitärer Ressourcen und Beratungsangebote

Bereits durchgeführte Gründungen befragter Mitarbeiterinnen und Mitarbeiter der Universität Bonn verliefen in 81% der Fälle ohne die Nutzung universitärer Ressourcen. Lediglich universitäre Räumlichkeiten wurden in wenigen Fällen (14%) genutzt. Bei den Befragten, die aktuell ein Unternehmen gründen oder sich dies für die Zukunft vorstellen können, ist die geplante Nutzung universitärer Ressourcen sehr viel stärker ausgeprägt. Nur 26% der Universitätsmitarbeiter mit Gründungsinteresse gaben an, kein Interesse an der Nutzung universitärer Ressourcen zu haben. Großes Interesse besteht vor allem an der Nutzung von Büroflächen (39%) und Büroausstattung (32%). Labore, Werkstätten und Gerätschaften würden 28% der potenziellen Gründer gerne nutzen.

Die befragten Spinoffs nutzten für die Gründung aus der Universität heraus im Gegensatz zu den befragten Universitätsmitarbeitern wesentlich häufiger universitäre Ressourcen. Nur in 10% der Fälle wurde hier auf die Nutzung universitärer Ressourcen verzichtet.

Zusätzlich wurde im Rahmen der Befragung erhoben, inwieweit die unterschiedlichen Beratungsangebote für Unternehmensgründungen von Seiten unterschiedlicher Anbieter bekannt sind und genutzt werden. Hierzu zählten insbesondere die Angebote der Universität Bonn und der Wirtschaftsförderung der Stadt Bonn. Gerade im Zusammenhang mit der Frage nach der Nutzung räumlicher Ressourcen fällt auf, dass die Beratung zum Inhouse-Modell der Universität wenig bekannt ist – nur 9% aller befragten Mitarbeiterinnen und Mitarbeiter haben schon einmal etwas hiervon gehört. Eher unbekannt ist ebenfalls der Stammtisch der Universitäts-Gründer (13% der Befragten). Das Seminarangebot im Rahmen des Programms IQU ist zwar 66% der Befragten bekannt, daran teilgenommen haben jedoch insgesamt nur 11% aller befragten Mitarbeiterinnen und Mitarbeiter. Sowohl der Nutzen der Teilnahme hinsichtlich der Vermittlung von Schlüsselkompetenzen als auch hinsichtlich der Gründungsvorbereitung wird von der überwiegenden Anzahl der Teilnehmer (70%) als hoch oder eher hoch eingestuft. Die Bekanntheit der verschiedenen Beratungsangebote der Universität ist in Abbildung 2 aufgeschlüsselt nach verschiedenen Gruppen dargestellt. Hier zeichnet sich vor allem ab, dass Gründerinnen und Gründer von Spinoff-Unternehmen wesentlich stärker von gründungsspezifischen Angeboten der Universität (Inhouse-Beratung und Stammtisch) Notiz nahmen als die übrigen Gruppen der Mitarbeiterinnen und Mitarbeiter.

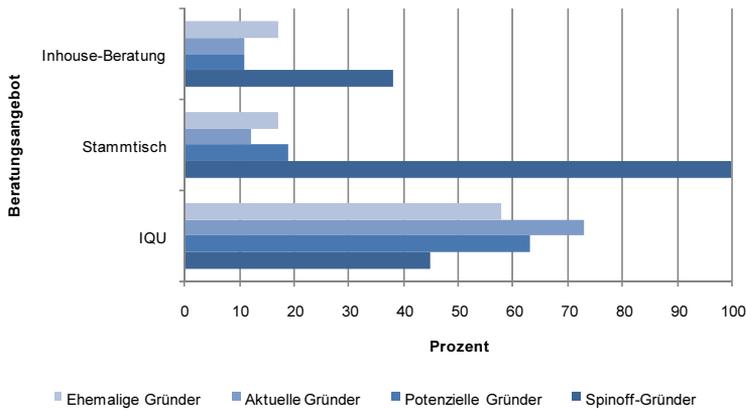


Abbildung 2: Bekanntheit des universitären Beratungsangebots

Außeruniversitär bietet vor allem das Amt für Wirtschaftsförderung der Stadt Bonn zahlreiche Beratungsangebote, deren Bekanntheit ebenso erfragt wurde. Als sehr bekannt erwies sich hier insgesamt die Fördermittelberatung und die Beratungsstelle »Frau und Beruf«. Den Gründungscoach kennen ebenfalls zahlreiche Befragte. Die Bekanntheit der Unternehmensnetzwerke »10points« und »40plus« sowie Genehmigungsmanagement und Immobilienservice ist deutlich geringer. Eine Darstellung der Bekanntheit der Beratungsangebote innerhalb der verschiedenen Gruppen findet sich in Abbildung 3.

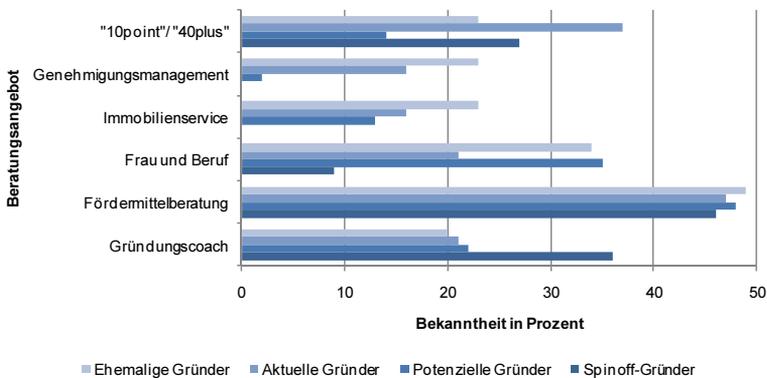


Abbildung 3: Bekanntheit des Beratungsangebots der Stadt Bonn

Verbesserungen im Bereich der Gründungsförderung werden von den Befragten in zwei Ansatzpunkten gesehen. Im Bereich der Beratung werden insbesondere zusätzliche Informationen zu kaufmännischen Grundlagen gefordert, die in der Phase des Unternehmensaufbaus für die Gründer wichtig sind. Als zweites wichtiges Thema ist die Risikobewertung einer Gründung für die potentiellen Unternehmer und Ausgründungen zu nennen. Hier sind Ansätze zur Abfederung des Gründungsrisikos erwünscht, die als Instrumente zur Gründungsförderung angesehen werden. Hierzu werden auch qualifizierte Informationen zur Unternehmensfinanzierung gezählt.

Forderung an die Universität

Sowohl nach Meinung der befragten Mitarbeiterinnen und Mitarbeiter als auch nach Meinung der Gründerinnen und Gründer von Spinoff-Unternehmen sollte die Vermittlung von gründungsspezifischem Know-how bereits im Studium verstärkt vermittelt werden. Die Zustimmung zu dieser Forderung ist hoch – unabhängig davon, ob Gründungserfahrungen vorliegen, oder nicht. Gerade vor dem Hintergrund der Tatsache, dass viele Gründungsideen bereits im Rahmen des Studiums entstehen, ist eine Förderung hier wünschenswert und sinnvoll. Auch die befragten Dekane unterstreichen diese Notwendigkeit unabhängig davon, ob aktuell Gründungsinteresse bzw. -potenzial in der entsprechenden Fakultät besteht oder nicht. Zukünftig wird der Bedarf nach alternativen Beschäftigungsfeldern und nicht zuletzt nach beruflicher Selbstständigkeit laut Aussage der Dekane unweigerlich größer, worauf es bereits in der Lehre vorzubereiten gilt.

Interesse an einem Gründerzentrum

Rückblickend besteht das größte Interesse an der Ansiedlung innerhalb eines Gründerzentrums unter Gründerinnen und Gründern von Spinoff-Unternehmen. Vergleichsweise gering schätzen dagegen befragte Mitarbeiterinnen und Mitarbeiter mit Gründungserfahrung ihr Interesse an einem Gründerzentrum zum Gründungszeitpunkt ein (33%). Abbildung 4 zeigt das Interesse der verschiedenen Gruppen an der Ansiedlung in einem Gründerzentrum.

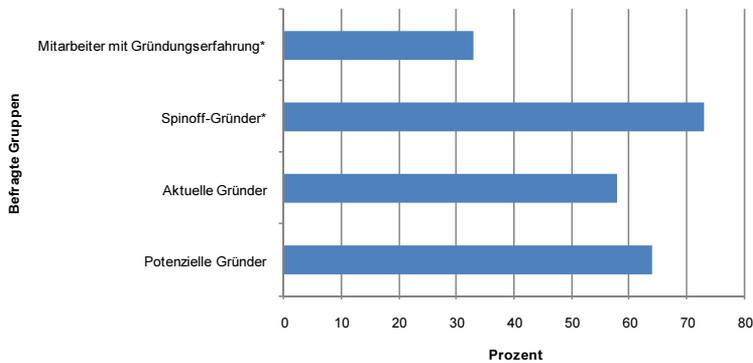


Abbildung 4: Prozentualer Anteil an Interessenten für eine Ansiedlung in einem Gründerzentrum

*Gründer von Spinoff-Unternehmen und ehemalige Gründer wurden gebeten rückblickend ihr Interesse zum Gründungszeitpunkt einzuschätzen.

15% der Mitarbeiterinnen und Mitarbeiter mit Gründungserfahrungen würden mit ihrem noch bestehenden Unternehmen gerne in ein Gründerzentrum umziehen, nur 8% machten bereits Erfahrungen mit der Ansiedlung innerhalb eines Gründerzentrums. Gründerinnen und Gründer von Spinoff-Unternehmen machten dagegen häufiger Erfahrung mit der Ansiedlung in einem Gründerzentrum (27%). Ebenfalls 27% der befragten Spinoffs würden mit dem bestehenden Unternehmen in ein neues Gründerzentrum umziehen.

Das optimale Gründerzentrum

Um für eine Ausgründung aus dem eigenen Arbeitsbereich interessant zu sein, sollte ein optimales Gründerzentrum nach Meinung der Befragten vor allem mit Büroflächen, Konferenzräumen und Laborflächen ausgestattet sein. 92% aller befragten Mitarbeiterinnen und Mitarbeiter der Universität bewerten Büroflächen als wichtig und eher wichtig. Besprechungsräume werden von 77% der Befragungsteilnehmer als wichtig und eher wichtig bewertet. Die Verfügbarkeit von Laborflächen wird von 63% für ein Gründerzentrum als wichtig oder eher wichtig bezeichnet. Produktionsflächen und Archivflächen spielen dagegen nach Meinung von nur unter 50% eine wichtige Rolle für die Attraktivität eines Gründerzentrums. Als weitere wichtige Standortfaktoren für ein Gründerzentrum werden eine gute Verkehrsanbindung und ein gutes Parkplatzangebot wahrgenommen.

Die Nähe eines Gründerzentrums zum eigenen Institut oder Arbeitsbereich wird als wenig wichtig eingeschätzt. So sehen nur 12% der Teilnehmer eine Distanz von mehr als 5 Minuten zwischen Gründerzentrum und eigenem Arbeitsbereich als inakzeptabel an. Die Mehrheit der Antwortenden bewertet

eine Distanz von bis zu 10 Autominuten vom eigenen Institut/Arbeitsbereich als optimal. Bezüglich der Büroinfrastruktur vermuten fast alle Befragten einen Vorteil, vor allem in bestehender Kommunikationsinfrastruktur und Geräteservice. Auch die Möglichkeit der Mobiliaranmietung wird von mehr als der Hälfte der Befragten als wichtig erachtet. Bezüglich möglicher Dienstleistungen werden IT-Service und Sekretariatsservice als besonders wichtig bewertet. Vor allem in der Vorgründungsphase und während der Gründung werden Beratungsdienstleistungen innerhalb eines Gründerzentrums von nahezu allen Befragten als elementar betrachtet.

Eine erhebliche Rolle für die Entscheidung zur Ansiedlung in einem Gründerzentrum spielt nach Meinung der Befragten vor allem die Möglichkeit zur Kooperation mit anderen dort angesiedelten Unternehmen. Zudem vermuten die Befragten unabhängig davon ob Gründungserfahrungen bestehen oder nicht, dass vor allem eine flexible und bedarfsgerechte Raumanmietung für die Entscheidung einer Ansiedlung in einem Gründerzentrum ausschlaggebend ist.

Blick auf vergangene Studien

Zur Bewertung der Ergebnisse der vorliegenden Befragung bietet sich der Vergleich mit einer ebenfalls vom Zentrum für Evaluation und Methoden und der Wirtschaftsförderung Bonn durchgeführten Ex-Ante Evaluation an. Hier wurden 330 Unternehmen in insgesamt 34 Gründerzentren zu Erfolgsfaktoren eines Gründerzentrums befragt. In dieser Studie wurde die tatsächliche Nutzung der Infrastruktur bzw. Ausstattung mit einer retrospektiven Einschätzung der gewünschten Ausstattung zum Gründungszeitpunkt verglichen.

Wie auch die Befragten in der vorliegenden Studie, bewerteten die Teilnehmer der Ex-Ante Evaluation die Möglichkeit zur flexiblen, bedarfsgerechten Flächenanmietung als wichtigen Grund für die Entscheidung zur Ansiedlung innerhalb eines Gründerzentrums. Im Gegensatz zu den Resultaten der vorliegenden Befragung sehen die Unternehmen der Ex-Ante Evaluation jedoch in Angeboten wie »bestehender Büroinfrastruktur« keinen wesentlichen Vorteil. Dieses Angebot erscheint zwar zunächst interessant, wird jedoch nach Aussage der Nutzer von Gründerzentren letztlich wenig genutzt.

Zusammenfassung und Fazit

Zusammenfassend lässt sich festhalten, dass das Gründerpotenzial der Universitätsmitarbeiter überraschend hoch ist. Das Interesse an der Ansiedlung in einem Gründerzentrum der Mitarbeiterinnen und Mitarbeiter, die sich mit dem Gedanken an eine Unternehmensgründung tragen, ist hoch. Voraussetzung für die Attraktivität eines solchen Gründerzentrums ist im Wesentlichen eine gute Anbindung und optimale Ausstattung sowie Nutzerfreundlichkeit. Bezüglich der Ausstattung ist jedoch mit Rückblick auf vergangene

Befragungen von einer Vollaussattung der Büroflächen mit Kommunikationsinfrastruktur abzuraten, da diese erfahrungsgemäß weniger genutzt wird, als von Gründern zunächst angedacht.

Das bestehende gründungsspezifische Beratungsangebot der Stadt Bonn und der Universität wird von den Nutzern überwiegend als nützlich bewertet – einige Themen, wie Risikominimierung und Finanzierung sollten jedoch nach Meinung der Befragten offensiver angegangen werden. Um bereits frühzeitig für das Thema Gründung zu sensibilisieren, sollte gerade in der Universität – die als Forschungseinrichtung und Innovationsstandort eine Geburtsstätte zahlreicher Gründungsideen darstellt – durch verstärkte Einbindung der Thematik in die Lehre auch das Potenzial zu deren praktischer Umsetzung gestärkt werden.

Uwe Feeser-Lichterfeld, Sandra Pietrangeli, Eva
Sindern, Uwe Kleinemas, Michael Fuchs & Georg
Rudinger
Lebensverlängerung und Verlangsamung des
menschlichen Alterns¹

Hintergrund und Fragestellung

Der Wunsch, dem körperlich-geistigen Altern trotzen und den altersbedingten Tod überwinden zu können, durchzieht die Menschheitsgeschichte. Er verbindet sich spätestens seit dem 17. Jahrhundert mit ähnlich gelagerten Utopien einer medizinischen Beherrschbarkeit der menschlichen Natur, wie sie im Gefolge und als Motor wachsender Kenntnisse und Fertigkeiten auf naturwissenschaftlich-technischem Gebiet generiert wurden und werden (Gordijn, 2004). Tatsächlich vermögen medizinische Interventionen inzwischen den Gesundheitsstatus älterer und alter Frauen und Männer so zu verbessern, dass deren Lebenserwartung kontinuierlich steigt. Ein Ende des Trends zur Langlebigkeit ist nicht abzusehen, die Annahme einer biologisch determinierten Obergrenze der menschlichen Lebenslänge erscheint zunehmend fragwürdig. Prognosen gehen dementsprechend von einem Anstieg der durchschnittlichen Lebenserwartung in Deutschland auf bis zu 95 Jahren bei den in 2050 Neugeborenen aus (Vaupel & Kistowski, 2007).

Szenarien, in denen der Seneszenzprozess verlangsamt, gestoppt oder gar rückgängig gemacht werden kann, erhalten durch Fortschritte bei der Entschlüsselung von Ursachen und Mechanismen des biologischen Alterns zusätzlich Aufwind. Der Optimierungsanspruch solcher biogerontologischer Projekte umfasst sowohl quantitative als auch qualitative Aspekte: Lebenslänge und Lebensqualität sollen durch die Therapie altersassoziierter Krankheiten und durch die Modulation des biologischen Alternsprozesses selbst gesteigert werden. Tierexperimentelle Erfolge deuten darauf hin, dass es möglicherweise nur eine Frage der Zeit ist, bis eine solche bislang nur schwer

1 Individuelle Einschätzung, gesellschaftliche Auswirkung, ethische Analyse und normative Beurteilung

Die Studie wurde im Rahmen der Förderinitiative »Geisteswissenschaften gestalten Zukunftsperspektiven« in den Jahren 2005 und 2006 durch das Ministerium für Innovation, Wissenschaft, Forschung und Technologie des Landes Nordrhein-Westfalen finanziert.

vorstellbare Verlängerung der Lebensspanne um zusätzliche Jahre in Gesundheit für den Menschen greifbar werden könnte (Arking, 2006; Austad, 2006; Holliday, 2006).

Wo die Möglichkeit einer direkt intendierten Verlangsamung des Seneszenzprozesses zunehmend in den Bereich einer realen Zieloption rückt, intensiviert sich die ethische Diskussion solcher biogerontologischer Ambitionen Feeser-Lichterfeld et al., 2007; Juengst & Binstock & Mehlman & Post & Whitehouse, 2003; Knell & Weber, 2009; Post & Binstock, 2004). Dabei stehen sich Argumente, welche die Natürlichkeit des menschlichen Lebenszyklus und die daraus abgeleitete Schutzwürdigkeit der Altersphase betonen, oftmals solchen Positionen diametral gegenüber, welche die aufweisbare Gestaltbarkeit des Lebenslaufs verteidigen wollen und biogerontologische Interventionen als Ausdruck einer präventiven bzw. therapeutischen Medizin interpretieren. Aus sozialetischer Perspektive werden entweder Allokationsprobleme aufgezeigt oder genau diese mit Verweis auf ohnehin ungleiche Lebensverhältnisse und Lebenschancen, unter denen Menschen in diachroner und synchroner Hinsicht leben müssen und die durch die diskutierten Techniken höchstens weiter divergieren, zu entkräften versucht.

Gerade die eigentümliche Polarisierung der Diskussion lässt es angeraten erscheinen, den ethischen Urteilsbildungsprozess um eine empirische Exploration individueller Einstellungen zu den projektierten biogerontologischen Eingriffsszenarien zu erweitern. Wichtigstes Ziel eines solchen Vorgehens ist es, den Wunsch nach Langlebigkeit sowie die potenzielle Nachfrage nach lebensverlängernden Maßnahmen abschätzen zu können, ihre Konsistenz zu prüfen und auf dieser Grundlage ihre ethische Akzeptabilität zu erörtern.

Studiendesign und Stichprobe

Die hier referierte Untersuchung, die in Kooperation von psychologischer Gerontologie und philosophischer Ethik konzipiert und durchgeführt wurde, war als Querschnittsstudie angelegt; im Zeitraum von Oktober 2005 bis April 2006 wurden 1.725 in Deutschland wohnhafte Frauen und Männer im Alter von 18 bis 90 Jahren mittels eines computerassistierten telefonischen Interviews (CATI) befragt. Die Stichprobe wurde zufällig gezogen und ist gemäß Mikrozensus repräsentativ für die Bevölkerung Deutschlands im Hinblick auf zentrale soziodemographische Aspekte wie Geschlecht und Alter. Tabelle 1.1 und Tabelle 1.2 stellen die relevanten soziodemographischen Daten dar.

Tab 1.1 Geschlechterverteilung in der Stichprobe

	<i>Häufigkeit</i>	<i>Prozent</i>
<i>männlich</i>	823	47,7
<i>weiblich</i>	902	52,3
<i>Gesamt</i>	1.725	100,0

Tab 1.2 Verteilung des Alters in der Stichprobe

		<i>Häufigkeit</i>	<i>Prozent</i>
<i>Gültig</i>	<i>18-24 Jahre</i>	191	11,1
	<i>25-34 Jahre</i>	260	15,1
	<i>35-44 Jahre</i>	425	24,6
	<i>45-54 Jahre</i>	330	19,1
	<i>55-64 Jahre</i>	269	15,6
	<i>älter als 65 Jahre</i>	240	13,9
	<i>Gesamt</i>	1.715	99,4
<i>Fehlend</i>		10	0,6

Als weitere soziodemographische Daten wurden der höchste Schulabschluss, der Familienstand und die Berufstätigkeit erfragt. Im Hinblick auf den jeweils höchsten Schulabschluss ist die Verteilung der Angaben wie folgt: Haupt-/Volksschulabschluss 20,2%, Realschulabschluss 30,0%, Fachabitur/Abitur 28,1%, Fachhochschul-/Hochschulabschluss 19,1%. Weitere 2,6% der Befragten haben keinen oder einen anderen Schulabschluss bzw. machen hierzu keine Angabe.

60,5% der Befragten sind verheiratet oder leben in einer festen Partnerschaft, 23,1% sind ledig, weitere 10,4% leben getrennt von ihrem Partner oder sind geschieden, 5,7% sind verwitwet. Von den Befragten, die nicht verheiratet sind oder in einer festen Partnerschaft leben, geben 43,8% an, in einer Partnerschaft zu leben, so dass insgesamt davon ausgegangen werden kann, dass 77,8% der Befragten in einer Partnerschaft leben.

61,6% der Befragten sind zum Zeitpunkt der Befragung berufstätig.

In Abhängigkeit von ihrem Antwortverhalten wurden den Befragten bis zu 100 Fragen gestellt, ein vollständig durchgeführtes Telefoninterview nahm im Mittel 20 Minuten in Anspruch. Erfragt wurden neben der Soziodemographie insbesondere Einstellungen zu einer Verlängerung der Lebensdauer und den hierfür in der Diskussion befindlichen Methoden, zu verschiedenen Aspekten des Alterns sowie zur Endlichkeit des Lebens. Weiterhin wurden der aktuelle Gesundheitsstatus, die Lebenszufriedenheit und die religiöse Orientierung erhoben. Zentral für die vorliegende Studie war die Frage nach dem persönlichen Wunschalter der Befragten unter den Bedingungen erfolversprechender biogerontologischer Eingriffe (»Stellen Sie sich vor, man könnte die Lebensdauer der Menschen tatsächlich nach Belieben verlängern, wie alt möchten Sie dann werden?«). Ziel der nachfolgenden Auswer-

tungen war es, die Befragten hinsichtlich dieses persönlichen Wunschalters und weiterer relevanter Merkmale zu typischen Gruppen zusammen zu fassen und die so identifizierten »Cluster« auf signifikante Unterschiede zu überprüfen.

Statistische Analysen und Ergebnisse

Die Angaben der Befragten zur Zielvariablen *Wunschalter* streuen extrem ($M = 94.715$ Jahre; $SD = 951.449$ Jahre; $Min = 30$ Jahre; $Max = 9.999.999$ Jahre). Der Median der Antwortverteilung liegt bei 90 Jahren.

Für die vorliegenden Analysen wurden Ausreißer in der Zielvariable 10%ig winsorisiert², wodurch eine Altersspanne von 80 bis 150 Jahre erzielt wurde ($M_{\text{winsorisiert}} = 98$ Jahre; $SD_{\text{winsorisiert}} = 20$ Jahre). Weiterhin wurden alle Items hinsichtlich ihrer Standardabweichung und ihrer Korrelation mit dem Kriterium *Wunschalter* untersucht. Items, die eine geringe Standardabweichung aufwiesen, wurden aus der Analyse ausgeschlossen, da davon ausgegangen werden musste, dass durch sie keine Unterschiede in der Zielvariablen erklärt werden können. Ebenfalls ausgeschlossen wurden Items, die mit der Zielvariablen nur gering korrelierten ($r < .2$) und für die somit kein statistisch bedeutsamer Zusammenhang zum Kriterium anzunehmen ist.

In dem so reduzierten Datensatz wurden anschließend fehlende Werte mittels Expectation-Maximization-Algorithmus (kurz EM-Algorithmus)³ ersetzt. Aufgrund faktorenanalytischer Ergebnisse konnten sechs Faktoren identifiziert werden. Die verbleibenden Items wurden aufgrund von inhaltlichen Überlegungen zu z-transformierten Skalenmittelwerten (Scores) zusammengefasst oder als Einzelitems in die Analysen aufgenommen (vgl. Abb. 1).

-
- 2 Die 10%ige Winsorisierung ersetzt jeweils die 10% größten und kleinsten Werte der Verteilung durch benachbarte Werte.
 - 3 Der EM-Algorithmus ist ein besonderes Schätzungsverfahren, das auf die Parameterschätzung bei unvollständigen Daten bzw. Missing Values spezialisiert ist. Hierbei wird sowohl das Antwortverhalten der Person als auch das Antwortverhalten aller Personen beim spezifischen Item berücksichtigt.

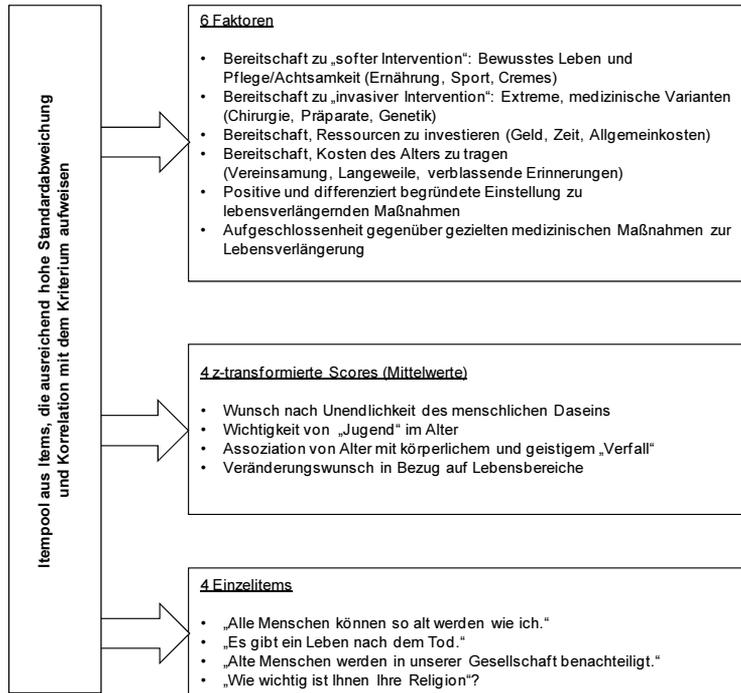
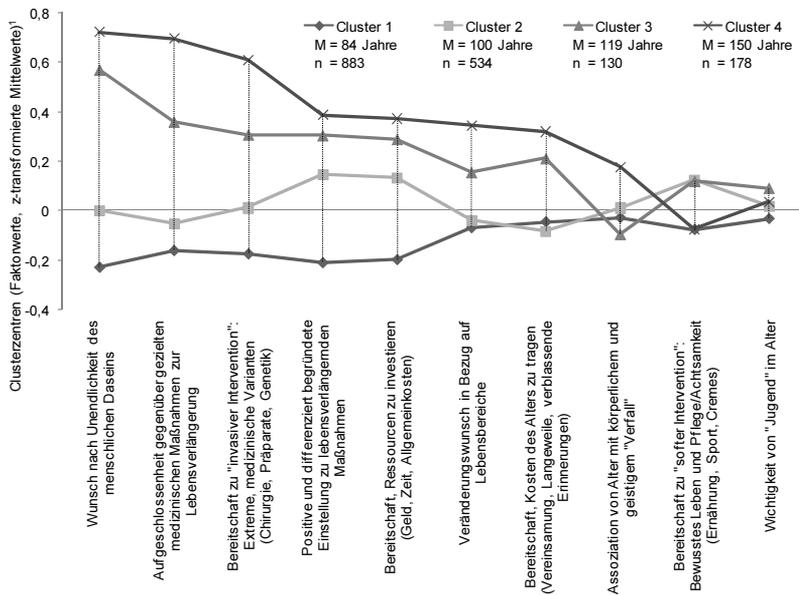


Abb.1 In die Analyse eingehende Faktoren, z-transformierte Scores und Einzelitems

Clusterzentrenanalyse

Zur Ermittlung von Gruppen, die sich hinsichtlich ihres Wunschalters und ihrer Einstellungen und Ansichten zum Altern möglichst ähnlich sind, wurde eine Clusterzentrenanalyse mit dem 10%ig winsorisierten Wunschalter, den Faktoren und den z-Scores als stabilen Dimensionen gerechnet. Hierdurch wurden vier Cluster identifiziert, die sich in erster Linie durch das Wunschalter beschreiben lassen (vgl. Abb. 3). Von den übrigen Variablen haben sich der Wunsch nach Unendlichkeit des menschlichen Daseins sowie die Aufgeschlossenheit gegenüber gezielten medizinischen Maßnahmen zur Lebensverlängerung und die Bereitschaft zu »invasiver Intervention« als besonders bedeutsam für die Clusterung erwiesen (vgl. Abb. 2).



¹ je höher der Wert, desto höher die Ausprägung auf der jeweiligen Dimension

Abb. 2 Clusterzentren der in die Clusterzentrenanalyse eingehenden Variablen (ohne Wunschalter)

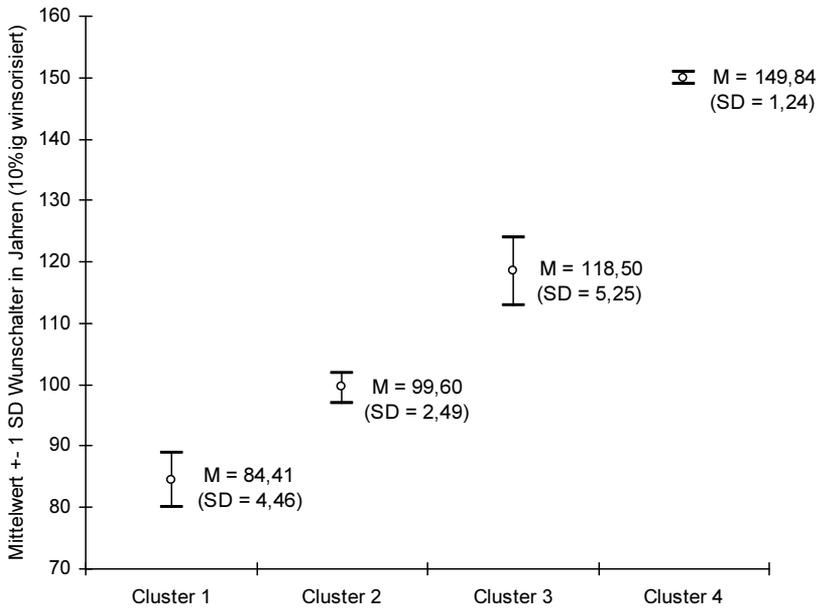


Abb. 3 Clusterzentren des 10%ig winsorisierten Wunschsalers: Mittelwert und Standardabweichung

Tabelle 2 zeigt die Distanz zwischen Clusterzentren der enggültigen Lösung. Es wird ersichtlich, dass die größte Distanz – und damit der größte Unterschied – zwischen den Clustern 1 und 4 besteht ($d_{\text{Cluster 1, Cluster 4}} = 65,452$), während die Cluster 1 und 2 bzw. die Cluster 2 und 3 einander ähnlich erscheinen ($d_{\text{Cluster 1, Cluster 2}} = 15,205$ und $d_{\text{Cluster 2, Cluster 3}} = 18,915$).

Tab. 2 Distanz zwischen Clusterzentren der endgültigen Lösung

Cluster	1	2	3	4
1		15,205	34,113	65,452
2	15,205		18,915	50,255
3	34,113	18,915		31,347
4	65,452	50,255	31,347	

Wichtig für die Interpretation der Ergebnisse ist die Verteilung der Befragten auf die Cluster: 51% (n = 883) der Befragten sind Cluster 1 zugehörig, 31% (n = 534) finden sich in Cluster 2. Cluster 4 ist mit einem Anteil von 10% (n = 178) vergleichsweise gering besetzt und lediglich 8% (n = 130) der Befragten gehören Cluster 3 an.

MANOVA

Um zu überprüfen, in welchen Merkmalen sich die identifizierten Cluster signifikant voneinander unterscheiden, wurde eine MANOVA mit der Clusterzugehörigkeit als unabhängige Variable und den Variablen der Clusteranalyse als abhängige Variablen gerechnet. Die in Tabelle 3 dargestellten multivariaten Tests verdeutlichen, dass sich zwischen den vier Clustern hochsignifikante Unterschiede ergeben (Pillai-Spur: $F = 78,137$; Hypothese $df = 33,000$; $p <= 0,000$).

Tab. 3 Multivariate Tests der MANOVA für Wunschalter

	Wert	F	Hypothese df	Fehler df	Sig.
<i>Effekt</i>					
<i>Pillai-Spur</i>	1,002	78,137	33,000	5139,000	,000
<i>Wilks-Lambda</i>	,033	335,569	33,000	5041,625	,000
<i>Hotelling-Spur</i>	28,616	1482,537	33,000	5129,000	,000
<i>Größte charakteristische Wurzel nach Roy</i>	28,579	4450,530 ^b	11,000	1713,000	,000

^bDie Statistik ist eine Obergrenze auf F, die eine Untergrenze auf dem Signifikanzniveau ergibt.

Die univariaten Post-Hoc-Tests nach Bonferroni und nach Tamhane ergaben dabei gleichermaßen, dass die häufigsten und größten signifikanten mittleren Differenzen zwischen den Clustern 1 und 4 bestehen. Dieser Befund deckt sich mit dem Ergebnis der Distanzmatrix der Clusterzentrenanalyse, die ebenfalls die größte Unterscheidung der Cluster 1 und 4 ergeben hatte. Für die Variable *Wunschalter* unterschieden sich alle vier Cluster wie erwartet für $p <= 0,01$ signifikant voneinander (vgl. Abb. 3). Jeweils drei der vier Cluster unterschieden sich auf dem 1%igen α -Niveau signifikant für die abhängigen Variablen *Wunsch nach Unendlichkeit des menschlichen Daseins* (kein signifikanter Unterschied zwischen Cluster 3 und 4), *Aufgeschlossenheit gegenüber gezielten medizinischen Maßnahmen zur Lebensverlängerung* (kein signifikanter Unterschied zwischen Cluster 1 und 2) und *Bereitschaft zu »invasiver Intervention«* (kein signifikanter Unterschied zwischen Cluster 2 und 3).

Die Ergebnisse der MANOVA sind konsistent mit denen der Clusterzentrenanalyse, die ebenfalls die größte Unterscheidung zwischen den Clustern hinsichtlich des persönlichen Wunschalters und der Einstellungen von Personen zur Begrenzung von menschlichem Leben sowie zur Aufgeschlossenheit gegenüber und zur Bereitschaft zur Anwendung von (auch

extremen) Maßnahmen zur Lebensverlängerung nahelegte (vgl. Abb. 2). Für die übrigen abhängigen Variablen unterschieden sich die Cluster weniger deutlich voneinander, für die Variable *Bereitschaft zu »softer Intervention«* konnte lediglich eine signifikante mittlere Differenz zwischen den Clustern 1 und 2 ermittelt werden. Für die Variablen *Wichtigkeit von »Jugend« im Alter* und *Assoziation von Alter mit körperlichem und geistigem »Verfall«* ergaben sich keine signifikanten Unterschiede zwischen den Clustern.

Einzelvergleiche

Abschließend wurde geprüft, welche bedeutsamen Unterschiede zwischen den Clustern in Bezug auf die vier in Abbildung 1 aufgeführten Einzelitems sowie auf Geschlecht und chronologisches Alter bestehen. Die vier Einzelitems lauten:

- (1) Alle Menschen können so alt werden wie ich.
- (2) Es gibt ein Leben nach dem Tod.
- (3) Alte Menschen werden in unserer Gesellschaft benachteiligt.
- (4) Wie wichtig ist Ihnen Religion?

Mit Ausnahme des Items *Wie wichtig ist Ihnen Religion?* ($n = 1.725$; $df = 3$; $F = 5,257$; $p = .001$) ergeben sich keine bedeutsamen Unterschiede in Bezug auf die Ausprägungen der Einzelitems zwischen den verschiedenen Clustern. Es zeigt sich, dass Personen, die ein hohes Wunschalter angeben, ein ausgeprägtes Bedürfnis nach der Unendlichkeit des menschlichen Daseins aufweisen sowie eine hohe Aufgeschlossenheit gegenüber gezielten medizinischen Maßnahmen und Eingriffen zur Lebensverlängerung zeigen, die Bedeutsamkeit der Religion signifikant niedriger einschätzen als Befragte, die in diesen Einstellungen eine geringe Ausprägung aufweisen und ein niedriges Wunschalter benennen (Mittlere Differenz = 0,2989; $p = .005$). Allerdings ist der Einfluss dieses Einzelitems auf die Varianzaufklärung mit einem Eta-Quadrat von 0,009 sehr niedrig.

Für die Variablen *Geschlecht* und *Alter* ergeben sich zwischen den Clustern jeweils signifikante Unterschiede, deren statistische Kennwerte in Tabelle 4 abgebildet sind.

Tab. 4 Statistische Kennwerte der Analysen für Geschlecht und Alter

Variable	N	df	Statistische Prüfgröße	p	Effektmaß
<i>Geschlecht</i>	1.725	3	$X^2 = 87,575$.000*	CC = .220
<i>Alter (insgesamt)</i>	1.715	3	$F = 14,396$.000*	$\eta^2 = 0,025$

* X^2 bzw. F sind auf dem ,01-Niveau signifikant.

Bei der Betrachtung der in Abbildung 4 dargestellten Verteilung von männlichen und weiblichen Befragten auf die Cluster fällt auf, dass Männer bedeutsam häufiger als erwartet in den Clustern 2, 3 und 4 vertreten sind. Sie geben also ein höheres Wunschalter an, zeigen einen ausgeprägteren Wunsch nach Unendlichkeit des Lebens und sind aufgeschlossener gegenüber medizinischen Maßnahmen und Eingriffen zur Lebensverlängerung. Frauen befinden sich vermehrt in Cluster 1 und weisen damit überzufällig häufig eine gegenteilige Ausprägung der genannten Merkmale auf.

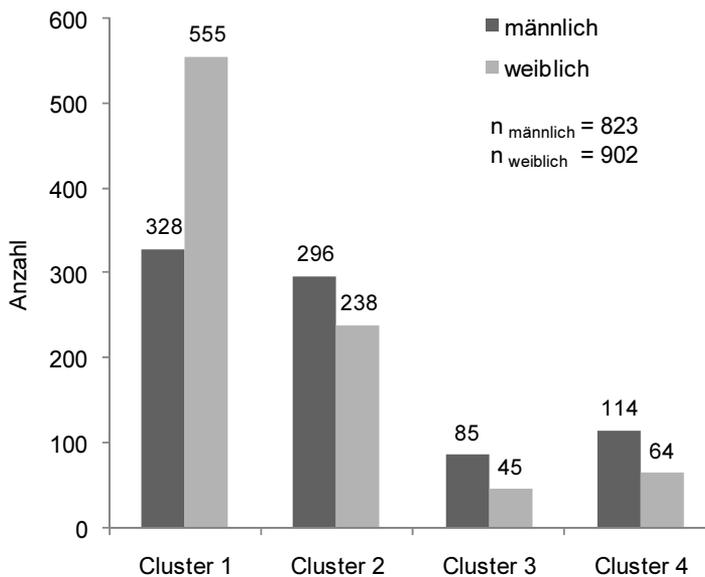


Abb. 4 Verteilung des Geschlechts auf die Cluster

Zudem wurde überprüft, ob der Verteilungsunterschied zwischen Männern und Frauen bestehen bleibt, wenn man das Alter der Befragten berücksichtigt. Dazu wurde die Stichprobe anhand des Alters in zwei Gruppen aufgeteilt: Personen, die bis zu 65 Jahre alt sind, und Personen, die älter als 65 Jahre sind. Ein nach dieser Aufteilung geschichteter Chi-Quadrat-Test zwischen Geschlecht und Clusterzugehörigkeit ergab, dass der zuvor beschriebene Effekt zwischen Männern und Frauen ausschließlich in der Gruppe der bis zu 65 Jahre alten Personen auftritt ($n = 1.506$; $df = 3$; $X^2 = 82,607$; $p = .000$; $CC = .228$).

Das chronologische Alter der Befragungsteilnehmer unterscheidet sich signifikant zwischen Cluster 1 und Cluster 4 ($M_{\text{Cluster1}} = 47$ Jahre; $M_{\text{Cluster4}} = 39$ Jahre; Mittlere Differenz = $8,3113$; $p = .000$) sowie zwischen Cluster 2 und Cluster 4 ($M_{\text{Cluster2}} = 45$ Jahre; $M_{\text{Cluster4}} = 39$ Jahre; Mittlere Differenz = $6,5343$;

$p = .000$). Dies bedeutet, dass Personen, die ein niedrigeres Wunschalter angeben, ein geringes Bedürfnis nach der Unendlichkeit des menschlichen Daseins aufweisen sowie gezielte medizinische Maßnahmen zur Lebensverlängerung und zu intensiven Interventionen ablehnen, bedeutsam älter als Befragte sind, die nach einem hohen Alter streben und bei denen die genannten Einstellungen entgegengesetzt ausgeprägt sind. Dieses Ergebnis bleibt unverändert, auch wenn das Geschlecht als zusätzlicher Faktor in die Analyse aufgenommen wird ($n = 1.715$; $df_{\text{Clusterzugehörigkeit}} = 3$; $F_{\text{Clusterzugehörigkeit}} = 14,155$; $p_{\text{Clusterzugehörigkeit}} = .000$; $df_{\text{Geschlecht}} = 1$; $F_{\text{Geschlecht}} = 0,557$; $p_{\text{Geschlecht}} = .456$; $df_{\text{Interaktion}} = 3$; $F_{\text{Interaktion}} = 0,474$; $p_{\text{Clusterzugehörigkeit}} = .700$).

Egoismus oder Solidarität – wie alt sollten Menschen im Allgemeinen werden?

Neben der Untersuchung der Cluster war der Zusammenhang zwischen dem individuellen Wunschalter und dem Wunschalter für die Allgemeinheit von besonderem Interesse. Deshalb wurden die Untersuchungsteilnehmer befragt, wie alt Menschen im Allgemeinen werden sollten, wenn man die Dauer des Lebens nach Belieben verlängern könnte. Wie bei der Angabe des individuellen Wunschalters wurden auch die Werte für das Wunschalter im Allgemeinen 10%ig winsorisiert und fehlende Werte durch den EM-Algorithmus ersetzt. Bei der Betrachtung der Verteilung der Werte des Wunschalters im Allgemeinen zeigen sich keine Unterschiede zum individuellen Wunschalter ($M_{\text{Allgemein}} = 98$ Jahre; $SD_{\text{Allgemein}} = 17$ Jahre). Auch der Zusammenhang zwischen individuellem Wunschalter und dem Wunschalter für Menschen im Allgemeinen ist stark positiv: Personen, die einen hohen Wert für ihr persönliches Wunschalter angeben, nennen auch für das Wunschalter im Allgemeinen hohe Werte, während Personen, die sich ein niedriges individuelles Wunschalter wünschen, auch für das Wunschalter für Menschen im Allgemeinen geringere Werte benennen ($r = .753$; $n = 1725$; $p = .000$).

Diskussion

Die Befragungsergebnisse zeigen deutlich, dass der im Titel dieses Aufsatzes unterstellte Wunsch nach einem möglichst unbegrenzt langen Leben von unserer Stichprobe weder unkritisch geteilt noch generell ins Reich unrealistischer Phantasien verwiesen wird. Stattdessen spiegeln die vier identifizierten Cluster differenzierte Standpunkte gegenüber den in Aussicht gestellten biogerontologischen Eingriffen wider. Diese reichen mit Blick auf die Lebenslänge von einem skeptisch-realistischen Beharren auf dem vertrauten Niveau (Cluster 1) über ein Setzen auf moderate Steigerungen (Cluster 2) und optimistisch anmutende Wünsche im Bereich bisheriger Rekordalter (Cluster 3) bis hin zu radikalen Perfektionierungsvisionen (Cluster 4). Dieser empirische Befund fördert den eingangs skizzierten und durch Polarisierung

gen zuweilen festgefahrenen ethischen Diskurs insofern, als dass das Szenario eines von Gebrechlichkeit und Endlichkeit soweit als möglich befreite Altern nicht nur im Hinblick auf seine biotechnische Realisierbarkeit und seine individual- wie kollektivethische Stimmigkeit analysiert werden kann, sondern zusätzlich unter Berücksichtigung des in der Bevölkerung tatsächlich verbreiteten Langlebigkeitswunsches. Die Tatsache, dass die große Mehrzahl der Befragten dabei – wenn überhaupt – für eine relativ maßvolle Verlängerung der Lebensspanne plädiert, vor allem aber die hohe Konsistenz von Wunschalterangaben einerseits und den Einstellungen zur Endlichkeit menschlichen Daseins und der jeweiligen Investmentbereitschaft zur Erreichung des eigenen Wunschalters andererseits, lässt das hier abgebildete Nachfragepotenzial umso ernsthafter erscheinen. Ethische Akzeptabilität kann solchen Positionen zumindest nicht ohne Weiteres abgesprochen werden. Denn zunächst erscheint es als durchaus plausibel, nicht nur das menschliche Leben als einen Wert, sondern auch die Verlängerung des Lebens als wertvoll anzusehen. Denjenigen, die sich ein im Vergleich mit bislang vertrauten Maßstäben deutlich längeres Leben wünschen, kann auch kein Egoismus unterstellt werden, da ihr persönliches Wunschalter zumeist mit den Antworten auf die Frage nach dem für die Allgemeinheit wünschenswerten Lebensalter übereinstimmt.

Unbeschadet der Beobachtung, wonach alle vier Cluster jeweils ein Muster aus Einstellungen, Erwartungen, Wünschen und Urteilen bilden, das prima facie als argumentativ konsistent und moralisch vertretbar angesehen werden kann, kann die spezifische Verteilung und das Zueinander evaluativer und normativer Aspekte in den Antworten unserer Befragungsteilnehmer nicht eindeutig bestimmt werden. Ob die das jeweilige Cluster prägende Haltung Ausdruck einer Zufriedenheit mit dem status quo, eines Argwohns gegenüber technischen Innovationspotenzialen, eines Mangels an Phantasie oder einer moralischen Skepsis im Blick auf die Folgen tiefgreifender technischer Selbstveränderung entspricht, lässt sich anhand unserer Daten nicht eindeutig darlegen. Diese interpretatorische Unsicherheit wurzelt nicht zuletzt in der Frage, inwiefern die Prämisse unseres Interviews – dass nämlich die thematisierten biogerontologischen Interventionen auch verfügbar und zudem erfolgversprechend seien – verstanden und von den Befragten bei Formulierung ihres Wunschalters beherzigt wurde. Dies vorausgesetzt, fordert der Umstand von vier Clustern mit einem je eigenen »considered judgment« dazu auf, den in Cluster 2 und Cluster 3 ausgedrückten »Mittelpositionen« mehr Gewicht in der ethischen Debatte einzuräumen bzw. eine ihrem Anliegen und ihrer Argumentationslogik entsprechende ethische Theorie zu formulieren.

Exemplarisch lässt sich diese Herausforderung einer »empirischen Ethik« konkretisieren: Gegner des Ziels einer verzögerten, gestoppten oder gar umgekehrten Seneszenz warnen dessen Anhänger vor dem in ihren Augen ver-

hängnisvollen Trugschluss, wonach mit Langlebigkeit auch eine Befreiung von alternstypischen Kontingenzindikatoren wie Gebrechlichkeit und Sterblichkeit einhergingen und sehen in der unterstellten Verdrängung der eigenen Endlichkeit ein kritikwürdiges Selbstmissverständnis. Im Blick auf unsere Studienergebnisse ist tatsächlich auffällig, dass für die Clusterung nach der Höhe des Wunschalers die persönliche Positionierung gegenüber der Endlichkeit bzw. Unendlichkeit des menschlichen Daseins leitend ist; hier replizieren die von uns Befragten in gewisser Weise das polarisierte Schema eingespielter ethischer Diskurse (einschließlich ihrer religiösen Bezugnahmen). Genau dieses vermag aber diskursförderliche Differenzierungen erfahren, wenn man auch die weiteren für die Clusterbildung maßgeblichen Aspekte und insbesondere die Einstellungen gegenüber lebensverlängernden Interventionstechniken beachtet. Hier wird deutlich, dass nicht allein anthropologische Grundpositionen für die registrierten Meinungsunterschiede verantwortlich sind, sondern zusätzlich die Frage, inwiefern ich Medizin und Biotechnologie die Ermöglichung von Langlebigkeit zutraue bzw. ob ich willens bin, mich derartigen Eingriffen zu unterziehen. In diesem Zusammenhang scheinen auch die beschriebenen Geschlechts- und Altersunterschiede interpretierbar zu sein.

Das differenzierte Meinungsbild, das unsere Befragung erbrachte, fordert nicht zuletzt zum Widerspruch gegenüber den Einseitigkeiten einer Anti-Aging-Bewegung auf, die versucht, das biologische Altern wie ein Krankheitsgeschehen zu interpretieren, dementsprechend zu therapieren und möglichst bereits im mittleren Erwachsenenalter zu stoppen. Hier wird insbesondere von Seiten der um individuelle und gesellschaftliche Durchsetzung eines differentiellen Alternsbildes ringenden Gerontologie, aber auch seitens der Verfechter einer Dignität der menschlichen Imperfektibilität zu hinterfragen sein, ob das in biogerontologischen Utopien verfolgte Ziel eines bedeutsam verlängerten Lebens und eines nach Möglichkeit nur noch chronologisch alternden Menschen eine Verschiebung unseres Selbstbildes beinhaltet, die als Verarmung zu deuten ist und die letztlich ein Ideal erzeugt, welches notwendig zu Frustrationen führt.

Literatur

- ARKING, R. (2006) *Biology of aging – Observations and principles (3rd ed.)*. New York: Oxford University Press
- AUSTAD, S.N. (2006). A biologist's perspective. Whence come we, where are we, whither go we? In D.J. Sheets, D.B. Bradley & J. Hendricks (Eds.), *Enduring questions in gerontology* (pp. 29-61). New York: Springer.
- FEESER-LICHTERFELD, U., FUCHS, M., ILLES, F., KLEINEMAS, U., KRAHN, B., PRELL, K., et al. (2007). Lebensverlängerung und Verlangsamung des menschlichen Alterns. Erträge eines interdisziplinären Forschungsprojekts. In L. Honnefelder & D. Sturma (Hrsg.), *Jahrbuch für Wissenschaft und Ethik* (Band 12, S. 219-254). Berlin: de Gruyter.
- GORDIEN, B. (2004). *Medizinische Utopien – Eine ethische Betrachtung*. Göttingen: Vandenhoeck & Ruprecht.
- HOLLIDAY, R. (2006). Aging is no longer an unsolved problem in biology. *Ann N Y Acad Sci*, 1067, 1-9.
- JUENGST, E.T., BINSTOCK, R.H., MEHLMAN, M.J., POST, S.G. & WHITEHOUSE, P. (2003). Biogerontology, »anti-aging medicine«, and the challenges of human enhancement. *Hastings Cent Rep*, 33 (4), 21-30.
- KNELL, S., WEBER, M. (Hrsg.). (2009). *Länger leben? Philosophische und biowissenschaftliche Perspektiven*. Frankfurt a.M.: Suhrkamp.
- POST, S.G. & BINSTOCK, R.H. (Eds.). (2004). *The fountain of youth. Cultural, scientific, and ethical perspectives on a biomedical goal*. Oxford: Oxford University Press.
- VAUPEL, J.W., KISTOWSKI VON, K.G. (2007). Die Plastizität menschlicher Lebenserwartung und ihre Konsequenzen. In P. Gruss (Hrsg.), *Die Zukunft des Alterns. Die Antwort der Wissenschaft* (S. 51-78). München: Beck.

Teil IV

Methoden und Statistik

Es werden drei Beispiele für Methodenentwicklung bzw. von statistischen Problemen im Anwendungskontext gegeben. Vor allem bei Umfragen liegen häufig selektive Stichproben mit zahlreichen fehlenden Werten vor, deren Auswertung mit gängigen statistischen Verfahren nur eingeschränkt möglich ist. In einem Artikel zum Radiohörverhalten (*»Zur Güte von Hörererhebungen in der Radioforschung¹«*) werden die Vor- und Nachteile längs- und querschnittlicher Erhebungen diskutiert sowie eine Untersuchung zur Übereinstimmung automatischer und gedächtnisbasierter Erfassung des Einschaltverhaltens berichtet. Verschärft tritt das Problem der Selektivität bei Panelanalysen, z.B. *Zukunftspanel* (IW Köln) und *Längsschnittstudien* auf.

¹ *Kursive Passagen* benennen bzw. umschreiben die Themen der Artikel

Norbert Hilger & Thomas Krüger

Zur Güte von Hörererhebungen in der Radioforschung¹

Zusammenfassung

Unter der Annahme, dass Quer- und Längsschnitte zufällig und mit Zurücklegen aus einer Grundgesamtheit gezogen werden, die hinsichtlich ihrer Kovarianzmatrix den media.lysis-Paneldaten des Jahres 2008 zur Radioforschung in Berlin entspricht, werden Vergleiche der Schätzgenauigkeiten beider Designs (einmalige Erhebung vs. Panel) zu den Kennwerten Hördauer und Nettoreichweite u.a. auf Jahresebene sowie für den Mittelwertsunterschied zweier aufeinander folgender Wochen durchgeführt. Ergebnisse sind, dass ein Panel mit täglicher Wiederholungsmessung zur Schätzung von Jahreswerten ungefähr 30% der Personenanzahl eines Querschnitts besitzen sollte, um eine ähnliche Genauigkeit aufzuweisen. Beim Wochenvergleich mit zweimal sechs Messzeitpunkten erreicht ein Panel bereits bei einer prozentualen Größe von 15% die Genauigkeit eines Trackings (zweimaliger Querschnitt).

Ferner wurde eine CATI-Befragung mit den media.lysis-Panelteilnehmern zum Hörverhalten des vorherigen Tages durchgeführt. Im Vergleich mit den media.lysis-Messwerten zeigt sich, dass die technisch erfassten Paneldaten kaum anhand einer Befragung gewonnen werden können. Die weitgehende Fehlerfreiheit der media.lysis-Messung vorausgesetzt, würde eine erinnerungsbasierte CATI-Befragung zu einer Überschätzung von kleinen Reichweiten (< 3.7%) führen.

1. Fragestellung

Die media.lysis GmbH erforscht im Raum Berlin das Radiohörverhalten im Rahmen eines Panels mit 197 Teilnehmern, wobei die Radionutzung auf Basis eines telemetrischen Erfassungssystems erfolgt, welches den Status der Empfangsgeräte minutengenau erfasst und in »Echtzeit« an das media.lysis-Rechenzentrum übermittelt. Der Untersuchungsansatz der media.lysis unterscheidet sich vom herkömmlichen Vorgehen der Radioforschung in Deutschland u.a. in zwei wesentlichen Punkten, nämlich dem längsschnittlichen statt wie üblich querschnittlichen Design sowie der telemetrischen Erfassung anstelle der telefonischen Befragung. Vor diesem Hintergrund ist die me-

¹ Der vorliegende Beitrag gibt Auszüge eines Gutachtens im Auftrage der media.lysis GmbH Magdeburg wieder.

dia.lysis mit den beiden folgenden Fragen an das Zentrum für Evaluation und Methoden (ZEM) der Universität Bonn herangetreten:

- (1) Wie viele Teilnehmer sollte ein Radiopanel haben, um ähnlich gute Ergebnisse wie ein Querschnitt zu liefern?
- (2) Wie gut können sich telefonisch befragte Personen an ihr Hörverhalten des Vortages erinnern?

Zur Beantwortung der ersten Frage werden zunächst Eigenschaften der media.lysis-Paneldaten des Jahres 2008 berichtet, aus denen sich unter bestimmten Annahmen ableiten lässt, in welchem Verhältnis die Anzahl der Panelteilnehmer zur Größe eines Querschnitts stehen sollte, um vergleichbare Genauigkeiten bei der Schätzung von Kennwerten der Radioforschung zu erzielen.

Um Anhaltspunkte zur Beantwortung der zweiten Frage zu erlangen, wurde eine CATI-Befragung durch das Telefonlabor des ZEM durchgeführt, in welcher die media.lysis-Panelteilnehmer nach dem Muster der herkömmlichen Radioforschung zu ihrem Hörverhalten des Vortages befragt wurden. Diese Befragungsdaten geben im Abgleich mit den telemetrischen Messungen der media.lysis einen Eindruck von der Güte der retrospektiven, erinnerungsbasierten Erfassung des Radiohörverhaltens.

2. Effizienz von Längsschnitt und Querschnitt

In der Radioforschung, wie sie in Deutschland z.B. von der Media-Analyse (MA) betrieben wird, werden typischerweise querschnittliche Untersuchungen durchgeführt. Hierbei wird das Hörverhalten von n Personen, verteilt über einen bestimmten Zeitraum (z.B. ein Kalenderjahr), einmalig erfasst, d.h. je Person wird das Hörverhalten von nur einem Tag ($k = 1$) abgefragt. Im Rahmen des media.lysis-Panels hingegen werden von m Personen Daten für jeweils jeden Tag eines Jahres, also für bis zu $k = 366$ Tage, kontinuierlich erfasst.

Wie sich nun die wiederholte Messung derselben Personen im Hinblick auf die Effizienz der Untersuchung auswirkt, hängt davon ab, wie sich das individuelle Radiohörverhalten über die Zeit entwickelt bzw. welche korrelative Struktur die Daten der Hörerforschung aufweisen: Wenn eine Person jeden Tag exakt dasselbe Hörverhalten zeigt, bringt eine wiederholte Messung über mehrere Tage keinen Informationsgewinn. Wenn hingegen eine Person jeden Tag ihre Sender ohne jegliche persönliche Präferenz zufällig wählt, kann die k -fache Messung dieser Person genauso viel Information enthalten wie die Messung von k unterschiedlichen Personen zu jeweils nur einem Zeitpunkt. Das sind die zwei Extremfälle einer völligen Abhängigkeit bzw. Unabhängigkeit des täglichen Radiohörverhaltens. Vermutlich wird sich das Verhalten der meisten Personen zwischen diesen Extremen bewegen,

wobei in der Regel von einer deutlichen, aber eben nicht perfekten Präferenzstruktur auszugehen ist.

Grundlage für die folgenden Vergleiche von Quer- und Längsschnittuntersuchungen ist der media.lysis-Paneldatensatz des Jahres 2008, der für 197 Personen das Einschaltverhalten von 358 Tagen enthält. Entsprechend der MA werden Messungen für Sonntage nicht berücksichtigt, womit 307 Tage verbleiben. Als Maß der Güte eines Untersuchungsergebnisses wird die Varianz des jeweiligen Kennwertes betrachtet und es wird gefragt, wie groß das m eines Längsschnitts sein sollte, damit ein Kennwert dieselbe Varianz besitzt wie die Schätzung auf Basis eines Querschnitts der Größe n .

Da sich die korrelativen Strukturen der media.lysis-Daten zwischen den Sendern unterscheiden, werden stellvertretend drei Sender analysiert, wobei »Sender 1« in Berlin relativ häufig, »Sender 2« mittel häufig und »Sender 3« eher selten gehört wird. Zudem beschränken wir uns auf zwei wesentliche Währungen der Radioforschung, nämlich die mittlere Hördauer in Minuten sowie die Nettoreichweite pro durchschnittlicher Stunde jeweils zwischen 6.00 und 18.00 Uhr. Da die meisten anderen Währungen auch auf diesen Mittelwerten basieren, lassen sich die Ergebnisse auch auf diese anderen Kennwerte übertragen. Im Folgenden werden zwei typische Situationen der Radioforschung untersucht: erstens die Bestimmung der genannten Kennwerte für ein ganzes Jahr und zweitens der Vergleich dieser Kennwerte zwischen zwei Wochen. Die zweite Situation wird zumeist mit einem Tracking, also auch anhand von Querschnitten beforscht.

2.1 Jahreswerte

Ziel der folgenden Berechnungen ist jeweils die Abschätzung der erforderlichen Größe m eines Panels mit k Messzeitpunkten, so dass ein Kennwert je Sender mit derselben Variabilität geschätzt werden kann wie anhand von Querschnitten der Größe n . Unter der Annahme, dass sowohl der Längsschnitt als auch der Querschnitt zufällig und mit Zurücklegen aus einer Grundgesamtheit mit den Eigenschaften der media.lysis-Paneldaten gezogen werden, lautet die Formel für die Varianz des Mittelwertes eines Querschnitts $Var_{\bar{x}}^{QS} = (\overline{Var}_{in} + Var_{zw})/n$, wobei $\overline{Var}_{in} = QS_{in}/(307 \cdot 196)$ die durchschnittliche Varianz innerhalb der Messzeitpunkte und $Var_{zw} = QS_{zw}/(306 \cdot 197)$ die Varianz zwischen den Mittelwerten der Messzeitpunkte des vorliegenden Paneldatensatzes sind. Die entsprechende Formel für den Längsschnitt lautet $Var_{\bar{x}}^{LS} = [\overline{Var}_{in} + (k-1)\overline{Kov}]/(km)$, wobei \overline{Kov} der Mittelwert der Kovarianzen aller Messzeitpunkte ist. Um nun abzuschätzen, wie groß m in Relation zu n gewählt werden muss, damit die Schätzgenauigkeiten der beiden Ansätze übereinstimmen, müssen diese beiden Formeln gleichgesetzt und umgeformt werden zu $m/n = [\overline{Var}_{in} + (k-1)\overline{Kov}] \cdot [k(Var_{in} +$

$Var_{zw})]^{-1}$; die rechte Seite dieser Gleichung entspricht dem Verhältnis der Varianzen $Var_{\bar{x}}^{LS} / Var_{\bar{x}}^{OS}$ für den Fall $m = n$.

In Tabelle 1 sind die Eigenschaften des media.lysis-Datensatzes 2008 zu den beiden Kennwerten Hördauer und Nettoreichweite für die drei ausgewählten Sender aufgeführt. Ferner sind für jeweils drei Konstellationen der Stichprobengrößen m und n je Sender die sich nach obigen Formeln ergebenden Verhältnisse der Varianzen (vorletzte Spalte) sowie der Streuungen (letzte Spalte) des Mittelwertes angegeben. Die unterstrichenen Werte geben jeweils den erforderlichen Stichprobenumfang eines Panels in Relation zur Größe eines gleichermaßen genauen Querschnitts an.

Tabelle 1. Varianzen und Kovarianzen von Hördauer und Nettoreichweite für drei Sender des Jahres 2008 ($k = 307$ Messzeitpunkte) sowie hieraus berechnete Verhältnisse der Varianzen und Standardfehler des Mittelwertes von Querschnitten der Größe n und Längsschnitten der Größe m in Abhängigkeit von der prozentualen Größe des Längsschnitts (in $100 \cdot \frac{m}{n} \%$)

		\overline{Var}_{in}	Var_{zw}	\overline{Kov}	$100 \cdot \frac{m}{n} \%$	$\frac{Var_{\bar{x}}^{LS}}{Var_{\bar{x}}^{OS}}$	$\frac{Se_{\bar{x}}^{LS}}{Se_{\bar{x}}^{OS}}$
Hördauer	Sender 1	6.171	0.034	1.609	100.0%	0.262	0.512
					<u>26.2%</u>	1.000	1.000
					10.0%	2.617	1.618
	Sender 2	4.159	0.022	0.787	100.0%	0.191	0.437
					<u>19.1%</u>	1.000	1.000
					10.0%	1.910	1.382
	Sender 3	1.979	0.007	0.667	100.0%	0.338	0.581
					<u>33.8%</u>	1.000	1.000
					10.0%	3.379	1.838
Netto-reichweite	Sender 1	77.589	0.372	30.091	100.0%	0.388	0.623
					<u>38.8%</u>	1.000	1.000
					10.0%	3.880	1.970
	Sender 2	58.118	0.344	16.664	100.0%	0.287	0.536
					<u>28.7%</u>	1.000	1.000
					10.0%	2.873	1.695
	Sender 3	18.405	0.079	5.656	100.0%	0.308	0.555
					<u>30.8%</u>	1.000	1.000
					10.0%	3.083	1.756

Sofern man also eine Grundgesamtheit mit den Eigenschaften der von der media.lysis beobachteten Daten unterstellt, ließe sich z.B. die durchschnittliche tägliche Hördauer für Sender 1 mit einem Längsschnitt der Größe m ähnlich genau schätzen wie mit einem Querschnitt der Größe n , wenn $100 \cdot \frac{m}{n} = 26\%$ ist. Für Sender 1 käme man im Längsschnitt gegenüber einem

Querschnitt also mit etwa einem Viertel der Personen zu einem vergleichbar genauen Ergebnis.

Abbildung 1 stellt die Relationen der Standardfehler von Längs- zu Querschnitt von Hördauer und Nettoreichweite für prozentuale Stichprobenumfänge des Längsschnitts zwischen 5% und 50% dar. Der Ordinate lässt sich (analog zur letzten Spalte in Tabelle 1) entnehmen, wie das Verhältnis der Schätzgenauigkeit von Längsschnitt zu Querschnitt in Einheiten des Standardfehlers ist. Anhand der vorliegenden Daten ist z.B. davon auszugehen, dass der Standardfehler der Hördauer auf Basis eines Panels mit einem Umfang von $100 \cdot \frac{m}{n} = 10\%$ für Sender 1 etwa 1.6 mal größer ist als auf Basis eines Querschnitts der Größe n . Die Schnittpunkte mit der Referenzlinie zeigen die erforderlichen Stichprobengrößen je Sender und Kennwert an, bei welchen von gleicher Genauigkeit des Längs- und Querschnitts auszugehen ist.

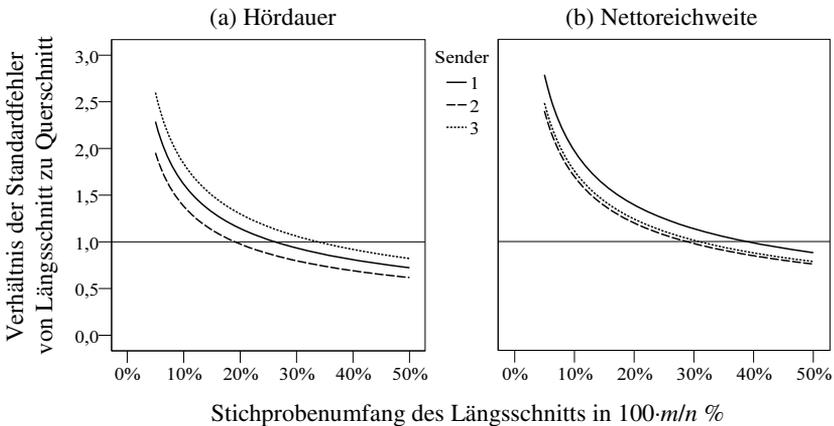


Abbildung 1. Vergleich der Schätzgenauigkeiten von (a) Hördauer und (b) Nettoreichweite (jeweils pro durchschnittlicher Stunde 6.00 bis 18.00 Uhr) für das Jahr 2008 eines Längsschnitts der Größe m mit $k = 307$ Messzeitpunkten und eines Querschnitts der Größe n in Abhängigkeit von der relativen Größe des Längsschnitts

Tabelle 2 führt zusammen mit den Punktschätzungen der für gleiche Genauigkeit von Längs- und Querschnitt erforderlichen Stichprobenumfänge Bootstrap-Konfidenzintervalle (Perzentil-Bootstrap mit 1000 Wiederholungstichproben) zu diesen Werten auf. Der Mittelwert der Punktschätzungen liegt bei 30%, und aufgrund der Konfidenzintervalle kann davon ausgegangen werden, dass der erforderliche Umfang für die meisten Kennwerte »tatsächlich« unterhalb von 40% liegt.

Tabelle 2. Für gleiche Schätzgenauigkeit von Längs- und Querschnitt erforderliche Stichprobenumfänge des Längsschnitts mit 95%-Konfidenzintervallen

		<i>untere Grenze</i>	$100 \cdot \frac{m}{n} \%$	<i>obere Grenze</i>
<i>Hördauer</i>	<i>Sender 1</i>	11%	26%	36%
	<i>Sender 2</i>	11%	19%	28%
	<i>Sender 3</i>	6%	34%	43%
<i>Nettoreichweite</i>	<i>Sender 1</i>	25%	39%	51%
	<i>Sender 2</i>	19%	29%	36%
	<i>Sender 3</i>	11%	31%	43%
<i>Mittelwert</i>		14%	30%	40%

Insgesamt lässt sich aufgrund der dargestellten Ergebnisse also sagen, dass ein Panel etwa einen Umfang von 30% der Größe eines Querschnitts besitzen muss, um zu ähnlich genauen Schätzungen von Jahreskennwerten der Radioforschung zu gelangen wie ein Querschnitt. Oder anders formuliert: Das vorgelegte media.lysis-Panel des Jahres 2008 mit $m = 197$ entspricht hinsichtlich seiner Effizienz zur Schätzung von Jahreswerten ungefähr einem Querschnitt der Größe $n = 650$. Je kleiner nun der Betrachtungszeitraum gewählt wird, desto größer wird der erforderliche Stichprobenumfang des Panels. Die den ausführlicher dargestellten Berechnungen zu Jahreswerten entsprechenden Ergebnisse (wiederum aggregiert über drei Sender und zwei Kennwerte) lauten auf Ebene eines Monats 47% und auf Ebene einer Woche 60%.

2.2 Wochenvergleiche

Im Folgenden wird die Varianz der Mittelwertsdifferenz für zwei aufeinander folgende Wochen betrachtet. Im Rahmen eines Trackings ergibt sich diese aus der Summe der jeweiligen Varianzen, $Var_{\bar{x}_1 - \bar{x}_2}^{OS} = Var_{\bar{x}_1}^{OS} + Var_{\bar{x}_2}^{OS}$, im Längsschnitt ist zudem die Kovarianz Kov_{12} zwischen den individuellen Mittelwerten der beiden Messzeiträume zu berücksichtigen: $Var_{\bar{x}_1 - \bar{x}_2}^{LS} = Var_{\bar{x}_1}^{LS} + Var_{\bar{x}_2}^{LS} - \frac{2 \cdot Kov_{12}}{m}$. Auch hier gilt, dass das Verhältnis der Varianzen für den Fall $m = n$ die erforderliche Relation von m zu n angibt, für welche die beiden Varianzen der Mittelwertsdifferenzen gleich groß ausfallen. Für das Jahr 2008 wurden nun für jeweils aufeinander folgende Wochen (1 vs. 2, 2 vs. 3, usw.) die Varianzen und Kovarianzen des media.lysis-Datensatzes berechnet (vgl. Tabelle 3) und Schätzwerte für die erforderlichen Stichprobengrößen des Panels wiederum in $100 \cdot \frac{m}{n} \%$ abgeleitet.

Tabelle 3. Auszüge der Ergebnisse zu Wochenvergleichen (Hördauer)

Sender	Woche	\overline{Var}_{in1}	Var_{zw1}	\overline{Kov}_1	\overline{Var}_{in2}	Var_{zw2}	\overline{Kov}_2	Kov_{12}	m/n
1	W1-W2	3.233	0.004	1.988	8.167	0.007	5.195	2.714	0.215
1	W2-W3	8.167	0.007	5.195	9.124	0.003	6.178	5.868	0.036
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
3	W49-W50	2.879	0.006	1.449	7.491	0.024	2.545	2.055	0.091

Abbildung 2 gibt die Verteilungen der für gleiche Effizienzen erforderlichen Stichprobengrößen des Panels in $100 \cdot \frac{m}{n} \%$ wieder. Jedem einzelnen Boxplot liegen 49 Werte auf Basis aufeinander folgender Wochen des Jahres 2008 zugrunde. Es zeigt sich, dass die erforderliche Stichprobengröße des Panels in aller Regel unterhalb von 20% liegt. In nur wenigen Ausnahmen ergeben sich (Schätz-)werte von 40% oder mehr.

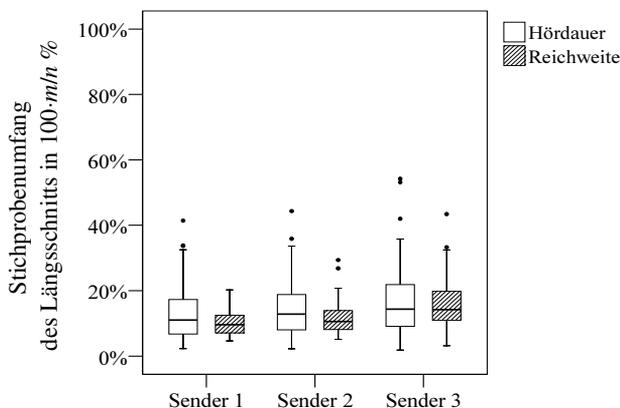


Abbildung 2. Verteilungen der erforderlichen Stichprobenumfänge von Längsschnitten für Vergleiche zwischen je zwei aufeinander folgenden Wochen (Montag bis Samstag) in % des Stichprobenumfangs von Querschnitten mit derselben Schätzgenauigkeit getrennt nach Sender und Kennwert (Hördauer und Nettoreichweite pro durchschnittlicher Stunde 6.00 bis 18.00 Uhr)

Der Mittelwert der 294 in Abbildung 2 dargestellten erforderlichen Stichprobenumfänge liegt über alle Sender und Kennwerte hinweg bei 15%. Vergleiche zwischen zwei aufeinander folgenden Wochen lassen sich auf Basis eines Panels mit $m = 197$ Teilnehmern also mit ähnlicher Genauigkeit durchführen wie anhand eines Trackings der Größe $n = 1300$ (d.h. zwei Querschnitte mit insgesamt 2600 Personen). Eine analoge Auswertung wurde auch für aufeinander folgende Monate durchgeführt, im Mittel liegt der erforderliche Stichprobenumfang dann bei 9%.

3. Überprüfung der Qualität der Erinnerungsleistung

Bei den bisher dargestellten Analysen wurde davon ausgegangen, dass sich die beiden Verfahren weder hinsichtlich der Gültigkeit noch der Genauigkeit (der einzelnen Messung) unterscheiden. Es ist aber anzunehmen, dass bei einer unerwarteten Abfrage die Erinnerung an das Hörverhalten des vorherigen Tages insgesamt nicht perfekt ist. Wir haben dazu eine empirische Studie mit den Teilnehmern des media.lysis-Panels durchgeführt. Ziel war es, zu messen, wie gut die Teilnehmer sich bei einem nicht angekündigten Anruf an ihr Hörverhalten des vorangegangenen Tages erinnern können.

3.1 Vorgehen bei der telefonischen Befragung

Insgesamt konnten 146 von 197 Personen des Panels befragt werden. Die telefonische Befragung erfolgte durch das Telefonlabor des ZEM. Die Personen wurden gebeten, an einer kurzen Befragung zur Zufriedenheit mit dem media.lysis-Gerät teilzunehmen. Nach wenigen Fragen zu diesem Thema wurden analog zum Vorgehen bei der MA Tätigkeiten, Aufenthaltsorte und Radionutzung des vorherigen Tages in Viertelstundenschritten erfasst. Gegenüber der MA, die die Radionutzung über den ganzen Tag erfasst, beschränkte sich die Abfrage in dieser Studie auf zwei Stundenintervalle (8.00 bis 9.00 Uhr und 17.00 bis 18.00 Uhr). Ferner war die Befragung so aufgebaut, dass falls eine Person angab, in einer bestimmten Viertelstunde einen Radiosender gehört zu haben, nachgefragt wurde, ob sie dazu das media.lysis-Gerät genutzt habe.

3.2 Ergebnisse des Vergleichs von media.lysis- und Befragungsdaten

Eine Übereinstimmung zwischen den retrospektiven Daten und den Messdaten von media.lysis liegt zum einen dann vor, wenn die Angabe einer Person, Sender X zum Zeitpunkt Y mit dem media.lysis-Gerät gehört zu haben, sich in den Daten, die vom Gerät registriert wurden, findet. Zum anderen liegt eine Übereinstimmung dann vor, wenn eine Person nicht angab, einen Sender gehört zu haben, und die media.lysis-Daten diesen Sender auch nicht als eingeschaltet ausweisen. Ein Fehler 1. Art (»Falsche Erinnerung«) liegt vor, wenn eine Person angab, zum Zeitpunkt Y einen Sender X gehört zu haben, aber keine entsprechende Nutzung des Gerätes vorlag. Ein Fehler 2. Art (»Fehlende Erinnerung«) schließlich liegt vor, wenn das Messgerät auf Sender X eingeschaltet war, die Person dies aber für diesen Zeitpunkt nicht berichtete.

Da von den 146 Befragten nur insgesamt 33 Personen in den zwei abgefragten Stunden überhaupt das media.lysis-Gerät eingeschaltet hatten, ist ein Vergleich der Erhebungsmethoden je Sender nicht sinnvoll, obwohl davon auszugehen ist, dass die Erinnerungsleistung zwischen den Sendern verschie-

den ist. Um zu einer stabileren Einschätzung hinsichtlich der Übereinstimmung von media.lysis-Messung und Befragung zu gelangen, wurde über alle insgesamt 20 Sender, die zumindest einmal gehört wurden, aggregiert. Diese Auszählung wurde zunächst auf Viertelstundenebene durchgeführt.

Wie Tabelle 4 aufzeigt, ist die Übereinstimmung der beiden Verfahren auf Viertelstundenebene insgesamt gering. Ein Maß für die überzufällige Übereinstimmung zweier Verfahren ist der Koeffizient Kappa, der Werte zwischen -1 und $+1$ annehmen kann, wobei $\kappa = 0$ zufällige Übereinstimmung und $\kappa = 1$ perfekte Übereinstimmung bedeuten. Kappa lautet hier $\kappa = .26$. Mit einer Befragung lassen sich die Ergebnisse des media.lysis-Verfahrens auf Viertelstundenebene also kaum gewinnen.

Tabelle 4. Vergleich zwischen media.lysis-Daten und erinnerungsbasierter Auskunft der Panelteilnehmer zum Hörverhalten auf Viertelstundenebene für die Zeiträume 8.00 bis 9.00 Uhr und 17.00 bis 18.00 Uhr des Vortages ($N = 146$ Befragte, Angaben zu 20 Sendern)

		<i>media.lysis</i>		
		<i>Sender eingeschaltet</i>	<i>Sender nicht eingeschaltet</i>	<i>Summe</i>
<i>Befragung</i>	<i>Sender gehört</i>	41	197	238
	<i>Sender nicht gehört</i>	29	23360	23389
<i>Summe</i>		70	23557	23627

Da die Forderung nach Übereinstimmung auf Ebene der Viertelstunde eventuell zu streng ist und Reichweiten vielfach auf Stundenebene berechnet werden, haben wir auch aggregiert auf Stundenebene ausgewertet. Eine Übereinstimmung liegt nun vor, wenn zumindest in einer der jeweils vier Viertelstunden berichtet wurde, dass ein Sender X mit dem media.lysis-Gerät gehört wurde und auch zumindest in einer der vier Viertelstunden das Gerät entsprechendes aufgezeichnet hat. Tabelle 5 zeigt die Ergebnisse bei dieser aggregierten Auswertung. Auch auf Stundenebene fällt die Übereinstimmung der beiden Verfahren nicht hoch aus ($\kappa = .36$). Insbesondere fällt auf, dass das media.lysis-Gerät nur in 25% ($= 100\% \cdot 28/113$) der Fälle, in denen ein Befragter einen bestimmten Sender genannt hat, auch auf diesen Sender eingestellt war. U.E. lassen sich diese Ergebnisse nicht mit der im Gegensatz zur MA hier zusätzlich geforderten Erinnerungsleistung hinsichtlich des media.lysis-Gerätes erklären, da die Mehrzahl der Befragten das Gerät im Auto betreiben und sich auch keine Unterschiede zu den Befragten mit Geräten im Haushalt ergeben.

Tabelle 5. Vergleich zwischen media.lysis-Daten und erinnerungsbasierter Auskunft der Panelteilnehmer zum Hörverhalten auf Stundenebene für die Zeiträume 8.00 bis 9.00 Uhr und 17.00 bis 18.00 Uhr des Vortages ($N = 146$ Befragte, Angaben zu 20 Sendern)

		<i>media.lysis</i>		
		<i>Sender eingeschaltet</i>	<i>Sender nicht eingeschaltet</i>	<i>Summe</i>
<i>Befragung</i>	<i>Sender gehört</i>	28	85	113
	<i>Sender nicht gehört</i>	10	5717	5727
<i>Summe</i>		38	5802	5840

Betrachtet man die media.lysis-Daten als (weitgehend) fehlerfreie Werte, so lassen sich die Eigenschaften der Befragungsmethode mit den Gütemaßen Sensitivität und Spezifität beschreiben. Sensitivität (SE) ist hier die Wahrscheinlichkeit, dass ein eingeschalteter Sender als solcher durch die Befragung erkannt wird: $\hat{SE} = 28/38 = .74$; und Spezifität (SP) gibt die Wahrscheinlichkeit an, dass ein nicht eingeschalteter Sender auch nicht als gehört angegeben wird: $\hat{SP} = 5717/5802 = .99$. Abbildung 3 zeigt nun für zwei Befragungsmethoden, die jeweils eine Spezifität von $SP = .99$ und eine Sensitivität von $SE = .74$ bzw. $SE = .85$ (letzterer Wert ergibt sich aus der Obergrenze eines 95%-Konfidenzintervalls für SE auf Stundenebene) besitzen, wie die Befragungsergebnisse von den wahren Werten im Mittel abweichen.

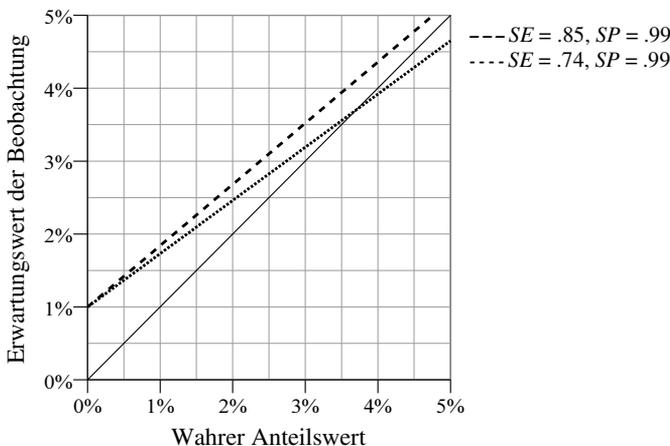


Abbildung 3. Beziehung zwischen wahren Anteilswert und Erwartungswert des beobachteten Wertes $[100\% \cdot p(SE + SP - 1) + 1 - SP]$ in Abhängigkeit von Sensitivität und Spezifität

Wenn z.B. ein Sender eine tatsächliche Reichweite von 1% hat, dann wäre zu erwarten, dass das Instrument mit $SE = .74$ eine Reichweite von 1.7% und das Instrument mit $SE = .85$ eine Reichweite von 1.8% erfasst. Für Reichweiten unterhalb von etwa 3.7% würden beide Verfahren zu einer Überschätzung des wahren Wertes neigen. Abbildung 3 verdeutlicht, dass kleine und sehr kleine Reichweiten nur mit einer Methode valide erfasst werden können, insbesondere deren Spezifität sehr hoch ist ($SP \geq .999$).

4. Diskussion

Sofern ein Längsschnitt zur Schätzung von »querschnittlichen Kennwerten« (wie z.B. Währungen der Radioforschung auf Jahresebene) verwendet wird, bei welchen die Abhängigkeiten über die Zeit keine ausdrückliche Berücksichtigung finden, zeigt sich bei der gegebenen Interkorrelation der media.ysis-Daten, dass eine dem Querschnitt gegenüber unverhältnismäßig hohe Anzahl von Messwerten erforderlich ist, um eine vergleichbare Genauigkeit der Schätzung zu erreichen. Um beispielsweise die Genauigkeit eines Querschnitts mit $n = 2000$ Messwerten (dies entspricht ungefähr dem Stichprobenumfang der Media-Analyse für Berlin) anhand eines Radiopanel mit 307 Messzeitpunkten zu erzielen, sind – wie in Abschnitt 2.1 gezeigt – 600 Teilnehmer und also 184200 Messwerte zu erfassen.

Wie in Abschnitt 2.2 für den einfachsten Fall der Mittelwertsdifferenz demonstriert wird, ergeben sich Vorteile des Panels dann, wenn die Abhängigkeiten zwischen den Beobachtungen berücksichtigt bzw. die Unterschiedlichkeit zwischen den Personen eliminiert werden. Auf Ebene eines Wochenvergleichs sind z.B. nur 15% der Personen je Querschnitt innerhalb eines Panels zu berücksichtigen. Im Vergleich zu einem Tracking mit einem üblichen Stichprobenumfang von z.B. $n_1 = n_2 = 300$, das sind insgesamt 600 Personen bzw. Messwerte, bedeutet dies, dass ein Panel mit 45 Personen und zweimal sechs Messzeitpunkten und also insgesamt 540 Messwerten zu einer vergleichbaren Schätzgenauigkeit führen kann. Das aktuelle media.ysis-Panel mit 197 Teilnehmern ist für Berlin also den herkömmlichen Trackings mit 200 bis 400 Personen je Querschnitt im Hinblick auf die Differenzschätzung deutlich überlegen. Doch auch hinsichtlich der Schätzungen pro Zeitintervall (Woche oder Monat) führt das media.ysis-Panel bei diesen Verhältnissen der Stichprobengrößen zu ähnlicher Genauigkeit (vgl. Abschnitt 2.1) wie ein Tracking.

Vorteile eines Panels ergeben sich also bei der Untersuchung von zeitlichen Effekten. Zudem sei hier auf die eigentlichen Besonderheiten eines Panels, nämlich die Möglichkeiten zur Beschreibung von Stabilität (z.B. der Senderpräferenzen) und individuellen Veränderungen (z.B. Hörerwanderung) hingewiesen. Anders als in der Fernsehforschung spielen diese Fragestellungen in der Radioforschung in Deutschland bisher ebenso kaum eine Rolle,

wie die Frage nach einer zum CATI-Befragungsansatz zusätzlichen oder alternativen technischen Erfassung des Radiohörverhaltens.

Um Anhaltspunkte für die Gültigkeit der erinnerungsbasierten Radioforschung zu erlangen, wurde anhand einer CATI-Umfrage unter den Panelteilnehmern nach dem Muster der MA die Übereinstimmung mit den technisch erfassten Messwerten der media.lysis berechnet. Es zeigt sich, dass die Übereinstimmung auch auf Ebene einer Stunde relativ gering ist (vgl. Tabelle 5). Insgesamt liegt die Zahl der erinnerten Sender etwa dreimal so hoch, wie die Zahl der von den media.lysis-Geräten aufgezeichneten Sender. Insgesamt liegen die Sensitivität der (hier realisierten) Befragungsmethode vermutlich unterhalb von .85 und die Spezifität bei maximal .99. Ein Verfahren mit diesen Eigenschaften würde insbesondere bei kleinen Reichweiten zu einer Überschätzung der tatsächlichen Reichweiten führen. Wesentliche Voraussetzung dieser Betrachtungen ist, dass die media.lysis-Daten fehlerfreie Angaben über den tatsächlichen Status der Empfangsgeräte machen. Hierzu lässt sich aufgrund unserer Untersuchungen keine Aussage machen.

Georg Rudinger

Analyse von Daten aus Längsschnittstudien: Eine Einführung

1. Was sind Längsschnittstudien?

Längsschnittstudien untersuchen ein- und dieselbe, aus einer wohl definierten Population gezogene Stichprobe von N Untersuchungseinheiten – i. d. R. Personen – zu den Zeitpunkten t_1, t_2, \dots, t_m , meist über längere Zeiträume, jedoch zu einer überschaubaren Anzahl von Gelegenheiten. Die Stichprobe bzw. die Elemente der Stichprobe verändern sich (sei es durch Intervention, durch systemischen Wandel des Kontexts oder auf natürliche Art, z. B. durch Alterung) in einem oder mehreren interessierenden Merkmalen V über den Untersuchungs-Zeitraum, so dass *intraindividuelle Vergleiche* möglich sind. Längsschnittstudien gibt es in verschiedenen Varianten: Prospektiver und retrospektiver Längsschnitt, *intensive longitudinal designs* (ILD), Panel-Studien, Trend-Studien, Zeitreihen, Ereignis-Analyse, Survival-Modelle, Kohorten-Studien (vgl. Diggle, Heagerty, Liang & Zeger, 2005; Frees, 2004; Singer & Willett, 2003; Taris, 2000; Walls & Schafer, 2006). Die verschiedenen Längsschnitt-Varianten lassen sich im dreidimensionalen Bezugsrahmen *Personen* (N) \times *Messzeitpunkte* (T) \times *Variablen* (V) positionieren. Sie unterscheiden sich in der Ausdehnung in einer oder mehrerer dieser Dimensionen: z. B. Panel-Studie – eher großes N ; Zeitreihe – eher kleines N , viele Zeitpunkte.

2. Personen \times Zeit \times Variablen: Selektivität

Der Raum *Personen \times Zeit \times Variablen* stellt das Universum dar, aus dem Stichproben von Beobachtungen für konkrete Studien gewonnen werden. Dabei kommt es zu Selektionseffekten. Besonders bzgl. der Personen (N) werden diese häufig diskutiert: Neben schwindendem Interesse können gerade bei Langzeitstudien auch äußere Gründe wie Wohnungs-, Schul- und Berufsartwechsel zum Ausscheiden (*drop out*) von Personen führen. Es kommt also i. d. R. zu Stichprobenminderung (*attrition*), die dann als systematischer Fehler zu werten ist, wenn sie kein Abbild eines zugrunde liegenden Populationsprozesses ist: Dies wäre z. B. der Fall, wenn in einer Längsschnittstudie mit Hochaltrigen die Sterberate in der Stichprobe geringer ist als in der Population. Selektionseffekte bzw. Minderungseffekte führen zu unvollständigen Datensätzen und damit zu Schwierigkeiten der Auswertung,

aber auch der (kausalen) Interpretation und der Generalisierung (Raudenbush, 2001). Selektivität ist in manchen Fällen *ex post* quantitativ zu erfassen und man kann ihr teilweise *a priori* aufgrund konkreter Annahmen oder Kenntnisse über die Mechanismen des Stichprobenschwunds durch gezielte kompensatorische Stichprobenziehung (*oversampling*) entgegen wirken. Nicht plan- und vorhersehbar ist, ob Personen einem bestimmten Messzeitpunkt fernbleiben und dann wieder teilnehmen oder innerhalb eines Messzeitpunktes einige Items nicht beantworten: Die Matrix *Personen x Zeitpunkte* hat bzgl. der Variablen also mehr (*unit nonresponse*) oder minder (*item nonresponse*) große »Löcher«. Wenn diese zufällig sind, handelt es sich um ein geringeres Problem. In der Regel aber sind diese *missings* systematischer Natur.

Dem Problem des Umgangs mit den verschiedenen Arten von *missing data* (MCAR – missing completely at random, MAR – missing at random, MNAR – missing not at random) widmet sich die Methodenforschung – besonders auch für den Längsschnitt – zunehmend (vgl. Graham, Taylor, Olchowski & Cumsille, 2006, die mit *missings* geradezu kreativ umgehen, und Reinecke, 2005, der einen konzisen Überblick über die verschiedenen Techniken des Ersetzens von *missing data* gibt).

3. Die Zeit-Dimension

Ein wesentliches Merkmal von Längsschnittstudien ist, dass die Daten eine Cluster-Struktur haben (Fitzmaurice, Laird & Ware, 2004), bzw. in verschiedene Ebenen gegliedert sind (Hedeker & Gibbons, 2004). Die jeweiligen Cluster bestehen aus den wiederholten Beobachtungen an je einem Individuum über die verschiedenen Zeitpunkte. Diese Cluster der wiederholten Messungen bilden dabei Ebene 1, während die Personen Ebene 2 darstellen. Längsschnittstudien haben (mindestens) diese Zwei-Ebenen-Struktur – oder anders ausgedrückt: Die wiederholten Beobachtungen sind in den Personen »genestet«. Vorteile der Messwiederholung sind darin zu sehen, dass die statistische Power in Längsschnittstudien größer ist und die Personen als Kontrolle ihrer selbst dienen. Die Messungen innerhalb des Clusters *Wiederholte Messungen über die Zeit* zeigen zumeist positive Korrelationen. Diese serielle Abhängigkeit der Beobachtungen erfordert differenzierte statistische Modelle, welche den Abhängigkeitsstrukturen zwischen den Wiederholungsmessungen Rechnung tragen.

Der Cluster- bzw. Ebenen-Gedanke lässt sich auf alle Dimensionen des $N \times T \times V$ -Bezugsrahmens ausdehnen: So gehören Personen Gruppierungen, Verbänden, Organisationen an wie Familien, Haushalten, Schulen, Nachbarschaften, Kohorten, Nationen etc., d. h. sind in ihnen genestet. Aus solchen den Individuen übergeordneten Beobachtungseinheiten können natürlich wiederum (Zufalls)-Stichproben gezogen werden.

Von zentralem Interesse ist bei Längsschnittstudien die Kovariation der untersuchten Merkmale (V) – der abhängigen Variablen (AV) – mit der Zeit (T). Nur dadurch ist es möglich, Prozesse zu beschreiben oder gar Kausalbeziehungen aufzudecken. Ein Problem allerdings ergibt sich, wenn nicht nur die AV mit der Zeit kovariiert, d. h. sich verändert, sondern auch Prädiktoren, Kontexte und Kovariaten dieses tun: Beispielsweise soll in einer Längsschnittstudie die Unfallgefährdung in Abhängigkeit vom Alterungsprozess untersucht werden; es ändert sich aber auch die Verkehrsdichte, die technische Ausstattung der Autos und die Aggressivität im Straßenverkehr.

In nicht-experimentellen Ansätzen wie Längsschnittstudien können die Einflussgrößen i. d. R. nicht willkürlich verändert werden. Es bleibt nur die Möglichkeit, *Veränderung* z. B. mithilfe von Kovarianz, Korrelation, Regression oder Übergangswahrscheinlichkeiten zu beschreiben und diesen Veränderungen vorausgehende Bedingungen zu identifizieren, um kausalen Beziehungen auf die Spur zu kommen. Zu diesem Zweck ist die Gerichtetheit der Zeit bei der Wahl des Designs, bei Modellbildung und bei der statistischen Analyse explizit zu berücksichtigen. Zeitliche Ordnung alleine ist allerdings keine hinreichende Bedingung für Kausalschlüsse (vgl. Taris, 2000). Korrelationen können auf nicht kontrollierbare Einflüsse von Dritt- und Kontextvariablen zurückgehen.

Bei Planung einer Längsschnittstudie gilt es, den Umfang und die Struktur der Zeit-Stichprobe – gezogen aus dem Universum der möglichen Zeitpunkte – dem jeweiligen Problem angemessen festzulegen (Singer & Willett, 2003). Die Anzahl von Messungen und die Abstände zwischen Messzeitpunkten (*spacing* und *timing*) sind über eine *causal lag theory* zu begründen, die sich Fragen gegenüber sieht, wie: Zu welchem Zeitpunkt – nach einem Ereignis, einer Intervention – könnte ein Phänomen auftreten – man denke an Inkubationszeiten? Wie lange ist die Dauer eines Prozesses, die Verweildauer in einer (Entwicklungs-)Phase bis zum Übergang in die nächste? Welche Zeitspanne der Konsolidierung ist notwendig, damit die Wahrscheinlichkeit eines Rückfalls (wieder) im »Normalbereich« liegt? (vgl. Collins, 2006). Ob Prozesse mit ihren unterschiedlichen Rhythmen und Geschwindigkeiten valide abgebildet werden können, hängt also von Lage und Anzahl der Messzeitpunkte ab. So ist es wenig zweckmäßig z.B. das Einkommen oder psychometrische Intelligenz für Veränderungsanalysen im Tagesabstand zu erfassen. Die Festlegung gleicher Abstände zwischen den Messzeitpunkten erscheint (z. B. angesichts theoriegeleiteter Erwägungen über den Einsatz der Wirkung einer Therapie und deren Nachhaltigkeit) oft willkürlich und wird i. d. R. eher aus forschungspraktischen Gründen vorgenommen.

4. Intraindividuelle Veränderung: Beschreibung und Erklärung

Nur Längsschnittstudien mit ihren – über wohl definierte Zeiträume sorgfältig geplanten – wiederholten Messungen ermöglichen direkte Aussagen über intraindividuelle Veränderungen. Dabei ist intraindividuelle *Veränderung* tunlichst von intraindividueller *Variabilität* zu unterscheiden (Nesselroade, 1991). Erstere ist systematischer Natur und impliziert langfristige, dauerhafte und unumkehrbare Veränderung (z.B. Intelligenzveränderungen über den Lebenslauf, also Veränderungen von *traits*), letztere impliziert reversible kurzfristige Veränderungen (Launen, Stimmungen, Emotionen – *states*). Beobachtete Unterschiede zwischen Personen zu einem bestimmten Zeitpunkt reflektieren neben diesen beiden Einflüssen aber auch stabile interindividuelle Differenzen (zwischen *traits*) sowie den Fehler der Messung.

Beschreibung, Erklärung und Vorhersage intraindividueller Veränderungen sowie interindividueller Differenzen und Ähnlichkeiten in intraindividuellen Veränderungen stellen nun die Hauptaufgabe für Längsschnittstudien dar. Intraindividuelle Veränderungen zeigen sich in verschiedener – quantitativ aber auch qualitativ modellierbarer – Form: Wachstum oder Minderung einer Fähigkeit (z.B. Intelligenz) werden beschrieben durch Mittelwerte, (Stabilität von) Verlaufskurven etc. Veränderung von Strukturen als Veränderung der Vernetzungsdichte oder -stärke (d.h. Differenzierung, Integration) wird beschrieben durch Maße von Zusammenhängen, wie Kovarianzen, Korrelationen oder als komplexer Aufbauprozess im Sinne von Sequenzen, Stufen, Hierarchien. Letztlich geht es darum, die Bedingungen solcher Veränderungen für Individuen oder Gruppen zu identifizieren (Konditionalanalyse), aber auch durch Modellierung von Wirkungszusammenhängen die Ursachen für Veränderungen zu ermitteln (Kausalanalyse).

5. Facetten von Konstanz und Wandel

An der Datensituation $T=2$ sollen einige grundlegende Facetten von Veränderung (und Stabilität) systematisch verdeutlicht werden. In Abbildung 1 werden die beobachteten Werte von jeweils einigen Personen zu zwei Zeitpunkten dargestellt.

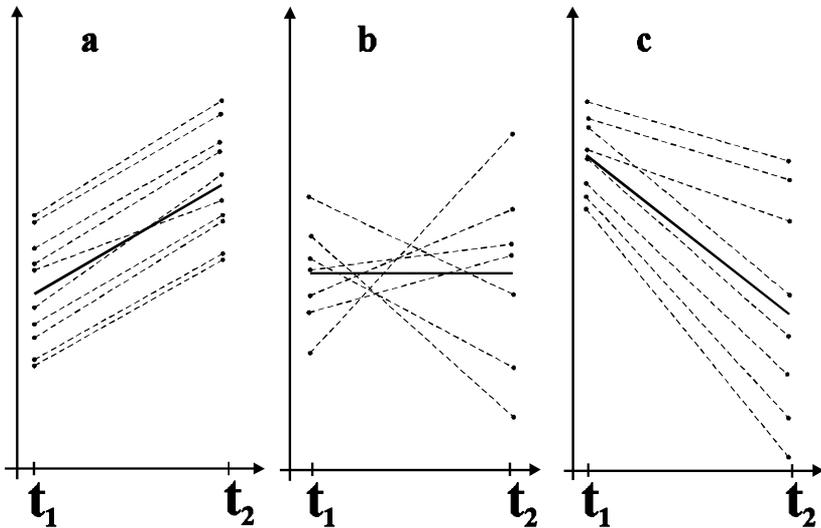


Abbildung 1: Längsschnittliches »Mini-Modell«: Eine Variable ($V=1$), zwei Messzeitpunkte ($T=2$).

Legende: Unterbrochene Linien: Individuelle Verläufe; Durchgezogene Linien: Mittelwertverläufe

Diese Abbildung zeigt prototypisch drei Muster (von vielen möglichen) individueller Verläufe (gestrichelt) über zwei Messzeitpunkte. Die durchgezogenen Linien verbinden die Mittelwerte zu t_1 und t_2 . Information über die Varianz der Messwertreihen ist der jeweiligen Spannweite der Messwerte zu t_1 und t_2 zu entnehmen. Änderungen in der Verteilung der Messwerte zu t_1 und t_2 können natürlich ebenfalls von Interesse sein, werden hier aber nicht betrachtet.

Typ a ist durch steigende Mittelwerte, durch stabile Varianzen und vor allem dadurch charakterisiert, dass die relativen Positionen der Personen fast unverändert bleiben, d. h. es gibt kaum interindividuelle Differenzen in intraindividuellen Veränderungen. Dies ist der Fall, wenn z. B. ein genereller, für alle Individuen gleichermaßen positiver Trainingseffekt vorliegt. Die – dadurch hier hohe positive – Korrelation der beiden Messwertreihen (r_{12}) wird auch als Test-Retest-Korrelation bezeichnet. Die Korrelation zwischen den Erhebungen zu zwei Messzeitpunkten (z. B. Intelligenz in der Schulzeit und im Erwachsenenalter) wird im Längsschnittkontext jedoch eher als Indikator für Stabilität betrachtet. Damit wird die Frage beantwortet: Wie stabil sind interindividuelle Differenzen über die Zeit? Es geht also nicht primär um die

Stabilität oder gar Art und Form der intraindividuellen Veränderung als Charakteristik eines einzelnen Individuums.

Typ b ist charakterisiert durch eine eher negative Korrelation (r_{12}), denn die Positionen der Personen wechseln mehrheitlich von Zeitpunkt zu Zeitpunkt gegenläufig. Die Varianzen indizieren wachsende interindividuelle Differenzen, dabei bleiben die Mittelwerte unverändert. Beispiel: Die Ausprägung der Lebenszufriedenheit zu zwei weit auseinander liegenden Zeitpunkten, zwischen denen es viele personenspezifische und unterschiedlich bewertete Lebensereignisse gegeben hat (z.B. Gewinn neuer Freiheiten durch Pensionierung, Veränderung der Wohnsituation, Verlust des Partners oder neue Partnerschaft).

Typ c ist bei fallenden Mittelwerten charakterisiert durch wachsende Varianzen und vor allem durch Beibehaltung der relativen Positionen fast aller Personen über die Messzeitpunkte. Wiederum ergibt sich also eine hohe positive Korrelation (r_{12}). Beispiel: Die Alltagskompetenz lässt mit steigendem Alter generell nach (fallende Mittelwerte); je niedriger allerdings das Ausgangsniveau einer Person schon ist, desto gravierender die Verluste – dies führt zu der Fächerform.

Die Statistiken (Mittelwert, Varianz und Korrelation - r_{12}) können also unabhängig voneinander über die Zeit variieren: Mittelwerte bleiben gleich, fallen oder steigen, Varianzen bleiben homogen oder werden heterogen, die Korrelationen (r_{12}) sind positiv oder negativ, hoch oder niedrig – jede Kombination ist möglich.

Bislang befanden wir uns auf der Beobachtungsebene: Ein Anwachsen der *beobachteten* Varianz zwischen Messzeitpunkten kann auf wachsende *wahre* Varianzen zurückgehen (d. h. »echte« Veränderung in der latenten Variablen, z. B. fächerförmiger *trait change*), könnte aber auch auf intraindividuelle Variabilität (*state*) oder einem Anwachsen des Messfehlers beruhen. Prinzipiell ist also die Beobachtungsebene von der latenten Ebene zu unterscheiden. Dies schließt die Varianzen auf beiden Ebenen ein, die Mittelwerte auf beiden Ebenen und die Kovarianzen und Korrelationen (r_{12}) auf manifester Ebene einerseits und latenter Ebene andererseits (ϱ_{12}). Je höher ϱ_{12} ist, desto ausgeprägter ist die Stabilität der *wahren* interindividuellen Differenzen über die Zeit, ganz im Sinne der vorherigen Stabilitätsdefinition. Es ist durchaus möglich, dass hinter jeder Facette der beobachteten Veränderung ein damit identischer oder davon verschiedener Veränderungsprozess in der latenten Welt liegt. Für dessen Identifikation spielt die Reliabilität eine Schlüsselrolle. So ist die beobachtete Korrelation über zwei Messzeitpunkte (r_{12}) das Produkt der Korrelation der latenten Variablen und der beiden Reliabilitätsindizes: $r_{12} = \rho_{12} \times \sqrt{rel_1 \times rel_2}$. (Anmerkung: Das Symbol »r« ist genau genommen ein Parameter, der die Korrelation der beobachteten Variablen in der Population bezeichnet). Reliabilität (*rel*) bezieht sich auf die Güte der Messung der latenten Variablen. Stabilität (ϱ_{12}) bezieht sich auf die Theorie über

den Prozess. Die beobachtete Korrelation ist eine Mischung aus beiden. Indiziert eine beobachtete niedrige Korrelation (r_{12}) über die Zeit nun niedrige Stabilität oder niedrige Reliabilität oder gar beides? Die Korrelation zwischen einer beobachteten Variablen zu zwei Messzeitpunkten kann nur dann als Grundlage für die Schätzung der Stabilität dienen, wenn die Reliabilitäten je Messzeitpunkt bekannt sind. Ist dies nicht der Fall, benötigt man zur separaten Schätzung von Reliabilitäten (*rel*) und ϱ_{12} mehrere Messungen pro *trait* und/oder Designs mit mehr als zwei Messzeitpunkten. Diese Schätzungen lassen sich dann mit Strukturgleichungsmodellen vornehmen.

6. Lineare Strukturgleichungsmodelle

Verknüpfungen zwischen der beobachteten und der latenten Welt können durch Strukturgleichungsmodelle (Structural Equation Modeling – SEM) hergestellt werden. Das Spektrum umfasst neben eher klassischen Modellen (s. dazu z. B. Bollen, 2002), bei denen Strukturfragen im Vordergrund stehen (s. Abschnitt 9), solche, die mit der Zeit variierende Variablen und Kovariaten behandeln und solche, die – ganz im beschriebenen Sinne – separate Hypothesen über sich verändernde Varianzen und Korrelationen – z. B. zur Schätzung von Reliabilität und Stabilität – testen und weitere, die auch Verläufe von Mittelwerten in die Modellierung einbeziehen, sog. LGC (Latent Growth Curve)-Modelle (Bollen & Curran, 2006; McArdle & Nesselroade, 2003; Plewis, 2005). Dabei lässt sich auch Gerichtetheit und Ordnung der Zeit explizit berücksichtigen (Curran & Bollen, 2001). Dies geschieht durch Modelle, in denen die latenten (d.h. reliablen) Anteile von Messwerten zum Messzeitpunkt t_i durch die latenten (reliablen) Anteile zum Messzeitpunkt t_{i-1} vorhergesagt werden, also durch den jeweils zeitlich unmittelbar vorausgehenden. Damit wird ein sog. autoregressiver Prozess erster Ordnung auf der Ebene der latenten Variablen modelliert. Der Vorteil der autoregressiven Modelle liegt darin, dass die Ordnung der Messzeitpunkte und die Gerichtetheit der Zeit essentielle Modellbestandteile sind. Ergebnis ist u. a. eine Kette von wahren Stabilitätskoeffizienten $\varrho_{t-1,t}$ zwischen direkt aufeinander folgenden Zeitpunkten. Über die Stabilitäten ($\varrho_{t-1,t}$) kann man ihrerseits Hypothesen formulieren, wie z. B., dass sie über die Zeit invariant bleiben ($\varrho_{t-1,t} = \varrho_{t,t+1}$; Stationarität) oder mit der Zeit größer werden ($\varrho_{t-1,t} < \varrho_{t,t+1}$); letzteres bedeutet, dass die interindividuellen Differenzen sich über die Zeit verfestigen.

Rosel und Plewis (2008) und Skrondal und Rabe-Hesketh (2004) geben gute Überblicke zu den vielfältigen Möglichkeiten und Anwendungen von SEM für Längsschnittstudien.

7. Ein Beispiel: Entwicklungsverläufe in den ersten beiden Lebensjahren

Längsschnittdaten über einen Zeitraum von zwei Jahren (1 bis 24 Monate) von neun Down-Syndrom-Kindern (aus einer Stichprobe von über 700) wer-

den in Abbildung 2 dargestellt (vgl. zu dieser Studie Rauh et al., 1991). Es handelt sich um die jeweils individuell diagnostizierten *Entwicklungs-Alter* (EA; Ordinate) im Bayley-Test (Mental Scale), abgetragen gegen die Zeitachse (chronologisches Alter – CA; Abszisse). Intraindividuelle Veränderungen liegen bei allen Kindern vor: Das Entwicklungs-Alter (EA) steigt mit dem chronologischen Alter (CA).

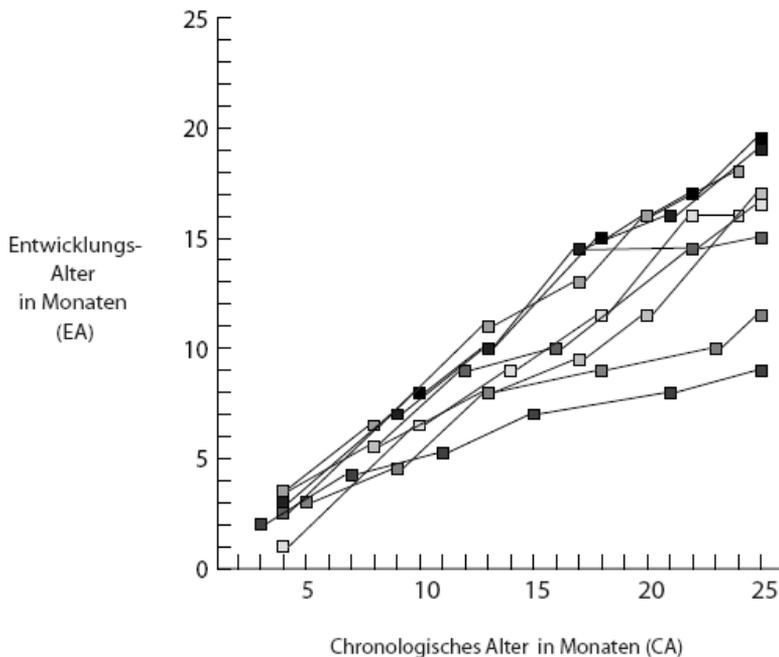


Abbildung 2: Längsschnittdaten von neun kanadischen Down-Syndrom-Kindern im Bayley-Test (Mental Scale) über 24 Monate.

Eine Analyse der Gesamtstichprobe mit einem Strukturgleichungsmodell, das auch Mittelwertsverläufe berücksichtigt (ein LGC-Modell also), zeitigt u. a. folgende Ergebnisse, die in der Abbildung 2 angedeutet sind: Die das Wachstum beschreibende Gleichung lautet: $EA = 0.54CA + 0.55$. Die Steigung (0.54) indiziert eine Verlangsamung der Entwicklung verglichen mit dem normalen Entwicklungstempo ($EA = 1.00 CA$). Die Standardabweichungen wachsen monoton über die Zeitpunkte (von 1.56 bis 4.96), d.h. die interindividuellen Differenzen werden größer. Die Stabilitäten (ρ) nehmen über die Zeit zu (von .47 bis .96), d. h. die relativen Positionen der Kinder verfestigen sich. Beides zusammen führt zur Fächerform. Die Reliabilitäten pro Messzeitpunkt sind moderat (im Schnitt .54) – dies ist aus der Abbildung natürlich nicht ersichtlich.

8. Mehrebenenmodelle – Hierarchische Lineare Modelle (HLM)

Eine weitere Frage könnte sich anschließen: Welches Populations-Wachstums-Modell (linear, kurvilinear, eines mit Sprüngen) hat die Stichprobendaten, genauer, die individuellen Verläufe generiert? Zur Beantwortung sollte in einem ersten Schritt durch jede der individuell beobachteten Wachstumskurven (Ebene 1, s. Abschnitt 3) eine Funktion gleichen Typs gelegt werden, die die empirischen Daten am besten repräsentiert (Cudeck & Harring, 2007). In unserem Beispiel gibt die lineare Regression jeweils die wahre intraindividuelle Veränderung adäquat wieder (within-subject model oder repeated-observations model). Die beobachteten individuellen Werte variieren natürlich jeweils um die individuelle Regressionsgerade (intraindividuelle Variabilität durch state und/oder sicher vor allem durch Messfehler). Die Regressionslinien haben pro Individuum unterschiedliche Steigungen (trait change) und Interzepte (trait differences), zeigen also (in diesen Statistiken) interindividuelle Varianz. Die Güte der Repräsentation der individuellen Kurven durch die »gemittelte« Regression (in unserem Falle: $EA = 0.54CA + 0.55$) wird durch das Ausmaß dieser Varianzen reflektiert (level-2 model oder person-level model). In unserem Beispiel sind diese Varianzkomponenten substantiell, d. h. ausgeprägte Heterogenität der Kinder liegt vor. Das legt die Vermutung nahe, dass es nicht nur diesen einen linearen Populationsprozess gibt, der für die Datenstruktur in der Stichprobe »verantwortlich« ist (s. Singer & Willett, 2003).

Eine unterschiedliche Prozesse bedingende Kontextvariable könnten die Lebensumstände der Kinder sein, z. B. Familienumgebung vs. Heimumgebung. Die Lebensumstände stellen also eine weitere Hierarchie-Ebene dar, innerhalb derer die Kinder »genestet« sind. Es ergebe sich für Kinder in Familien folgende Funktion: $EA = 0.65CA + 1.28$, für Kinder im Heim die Funktion: $EA = 0.38CA + 0.47$. Die beiden Gruppen von individuell beobachteten Verläufen sind in der Abbildung 2 veranschaulicht: Die unteren beiden mit moderaten Steigungen repräsentieren die Heimbedingung, die oberen sieben mit steileren Anstiegen die Familienbedingung. Diese beiden Gruppen scharen sich enger um die jeweilige spezifische Entwicklungsfunktion für »Heim« bzw. »Familie« als um die allgemeine Funktion für die gesamte Gruppe. Die Berücksichtigung zweier heterogener Prozesse (in diesem level-3 model) führt zu deutlich geringeren Varianzkomponenten als das allgemeine (level-2) Modell und »erklärt« somit die entsprechende Datenstruktur besser (s. dazu auch Bryk & Raudenbush, 2002; Frees, 2004; Keller, 2003; Raudenbush, 2001).

9. Strukturveränderungen

In der Studie von Rauh et al. (1991) wurde nicht nur mentale, sondern auch motorische Entwicklung untersucht. Eine weitere Frage könnte nun

sein: Wie sind mentaler und motorischer Bereich miteinander verknüpft? Wird deren Vernetzung über die Zeit enger oder lockert sich die Verbindung? Dies sind Fragen zur Struktur von Veränderungen und zu Veränderungen von Strukturen. Strukturelle Veränderungen würden sich in über die Zeit unterschiedlichen Korrelationen zwischen den Konstrukten zeigen bis hin zur sich verändernden Zahl von Faktoren (Differenzierung oder Integration). Solche Probleme lassen sich im Paradigma der faktoriellen Invarianz behandeln, in dem geprüft wird, ob »gleiche« Faktoren pro Messzeitpunkt, »ähnliche« Ladungen pro Faktor und Messzeitpunkt und »ähnliche« Beziehungen zwischen den Konstrukten pro Messzeitpunkt vorliegen (vgl. Cunningham, 1991). Damit sind die Bedeutungs-Invarianz von Konstrukten und die Validität der Messinstrumente über Zeit-Stichproben genauso angesprochen wie die über Stichproben von Personen, z. B. beim Vergleich von Kohorten (Labouvie, 1980). Nur bei Bedeutungs-invarianz/Identität eines Konstrukts über die jeweilige Dimension (T, N) sind Vergleiche individueller Wachstumskurven und deren Aggregation sinnvoll und aussagekräftig (Meredith & Horn, 2001).

Die meisten der mit struktureller Veränderung und Invarianz zusammenhängenden Fragestellungen lassen sich mit Linearen Strukturgleichungsmodellen behandeln (Bollen, 2002), zunehmend – wie gesagt – auch im Kontext zeitgerichteter Modelle (Rosel & Plewis, 2008).

10. Ausblick

Nur der Längsschnittansatz bietet die Möglichkeit, intraindividuelle Veränderungen und interindividuelle Differenzen und Ähnlichkeiten in den intraindividuellen Veränderungen zu beschreiben, deren Charakteristika und Determinanten zu ermitteln oder Veränderungen in Struktur oder Struktur der Veränderung zu untersuchen. Damit dies nach den Regeln der Kunst geschieht, muss »Zeit« vor allem bei Designs und Auswertungsmodellen, ihrem theoretischen und praktischen Status angemessen, Berücksichtigung finden. Die Verfahren zur Kausal- und Veränderungsanalyse (Lineare Strukturgleichungsmodelle – SEM, Latent Growth Curves – LGC -, Hierarchical Linear Modeling – HLM -, aber auch Latente Klassenanalyse) eröffnen die Möglichkeit, die Fragen adäquat zu beantworten, derentwegen Längsschnittstudien durchgeführt werden. Entsprechende Varianten von SEM (z. B. LGC) und HLM wachsen zusammen, bzw. können ineinander überführt werden. Ohne SEM und HLM ist heute keine Auswertung von Längsschnittstudien mehr möglich.

Eine abschließende Bemerkung zu weiterführender Literatur: Taris (2000) ist grundlegend gehalten und behandelt (deswegen) weder SEM noch HLM. Ersteres tut Chan (2003), zweites tun Hedeker und Gibbons (2004). Singer

und Willett (2003) geben einen durchaus anspruchsvollen, aber auch didaktisch sehr guten Gesamtüberblick.

Literatur

- BOLLEN, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, 53, 605-634.
- BOLLEN, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equation approach*. Hoboken, N.J.: Wiley.
- BRYK, A. S. & Raudenbush, S. W. (2002). *Hierarchical linear models: Application and data analysis methods*. Thousand Oaks: Sage.
- CHAN, D. (2003). Data analysis and modeling longitudinal processes. *Group & Organization Management*, 28, 341-365.
- COLLINS, L. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505-528.
- CUDECK, R. & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficients models. *Annual Review of Psychology*, 58, 615-637.
- CUNNINGHAM, W. R. (1991). Issues in factorial invariance. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 106-113). Washington, DC: American Psychological Association.
- CURRAN, P. J. & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp.105-135). Washington: American Psychological Association.
- DIGGLE, P. J., Heagerty, P., Liang, K.-Y. & Zeger, S. L. (2005), *Analysis of longitudinal data*. (2nd ed.). Oxford: Oxford University Press.
- FITZMAURICE, G. M., Laird, N. M. & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, N.J.: Wiley.
- FREES, E. W. (2004). *Longitudinal and panel data: Analysis and applications in the social sciences*. Cambridge: Cambridge University Press.
- GRAHAM, J. W., Taylor, B. J., Olchowski, A. E. & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, 11, 323-343.
- HEDEKER, D. & Gibbons, R. D. (2004). *Multilevel models for longitudinal data*. Hoboken, N.J.: Wiley.
- KELLER, F. (2003). Analyse von Längsschnittdaten: Auswertungsmöglichkeiten mit hierarchischen linearen Modellen. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 32, 51-61.
- LABOUVIE, E. W. (1980). Identity versus equivalence of psychological measures and constructs. In L.W. Poon (Ed.), *Aging in the 1980's: Selected contemporary*

- issues in the psychology of aging* (pp. 493-502). Washington, DC: American Psychological Association.
- MCARDLE, J. J. & Nesselroade, J. R. (2003). Growth curve analysis in contemporary psychological research. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology. Research methods in psychology* (pp. 447-480). Hoboken, N.J.: Wiley.
- MEREDITH, W. & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203-240). Washington: American Psychological Association.
- NESSELROADE, J. R. (1991). Interindividual differences in intraindividual changes. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 92-105). Washington, DC: American Psychological Association.
- PLEWIS, I. (2005). Modeling behaviour with multivariate multilevel growth curves. *Methodology, 1*, 71-80.
- RAUDENBUSH, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology, 52*, 501-525.
- RAUH, H., Rudinger, G., Bowman, T. G., Berry, P., Gunn, P. V. & Hayes, A. (1991). The development of Down Syndrome children. In M. Lamb & M. Keller (Eds.), *Infant development: Perspectives from German speaking countries* (pp. 329- 355). Hillsdale: Erlbaum.
- REINECKE, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München: Oldenbourg.
- ROSEL, J. & Plewis, I. (2008). Longitudinal analysis with structural equations. *Methodology, 4*, 37-50.
- SINGER, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. Oxford: Oxford University Press.
- SKRONDAL, A. & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.
- TARIS, T.W. (2000). *A primer in longitudinal data analysis*. London: Sage.
- WALLS, T. A. & Schafer, J. L. (Eds.). (2006). *Models for intensive longitudinal data*. New York, N.Y.: Oxford University Press

Simon Riek, Thomas Krüger & Georg Rudinger

Befragungsmethodische Aspekte des IW-Zukunftspanels¹

Die IW Consult GmbH, eine Tochter des Instituts der deutschen Wirtschaft Köln (IW Köln), baut seit 2005 ein Online-Befragungspanel von Geschäftsführern deutscher Industrie- sowie mit der Industrie verknüpfter Dienstleistungsunternehmen auf – das IW Zukunftspanel. An den drei Befragungswellen pro Jahr nehmen jeweils gut 3.000 Unternehmen teil. Jede Welle hat einen Themenschwerpunkt, bei dem es sich immer um Facetten des Strukturwandels der deutschen Industrie handelt, z.B. Fachkräftemangel oder Internationalisierung. Zudem werden umfangreiche Basisdaten, wie Mitarbeiterzahlen, Umsatz, Forschungsaktivitäten zu jedem Unternehmen erfasst. Das ZEM ist für die Durchführung der Online-Befragungen, die telefonische Erfassung der Email-Adressen der Geschäftsführer und die operative Betreuung (Versand Emails, Reminder etc.) zuständig.

Die Komplexität der Themen führt dazu, dass die durchschnittliche Bearbeitungsdauer des Fragebogens bis zu 30 Minuten umfassen kann. Die Teilnehmer werden nicht monetär entlohnt, aber erhalten für die Teilnahme exklusiv Ergebnisse der Befragungen. Die Länge des Fragebogens, die Komplexität einzelner Themenbereiche sowie die schwierige Erreichbarkeit der Zielgruppe, führen dazu, dass das Strukturpanel auch befragungsmethodisch sehr anspruchsvoll ist.

Unser Beitrag stellt die Probleme und unsere Lösungsansätze, die wir zusammen mit der IW Consult entwickelt haben, dar.

1. Einleitung

Ziel einer Umfrage ist es immer, eine möglichst exakte Beschreibung eines Zustands zu erhalten. An vielen Wahlabenden wird deutlich, dass es gar nicht so einfach ist, mithilfe einer Umfrage eine genaue Prognose abzuleiten. Leider ist die Größe einer Stichprobe – also die Anzahl der Befragten – alleine kein Garant für die Güte oder Repräsentativität einer Umfrage. Wichtiger dafür sind die Ziehung dieser Stichprobe und das Ziel, eine möglichst hohe Teilnahmebereitschaft und damit Ausschöpfung zu erreichen.

1 Der Beitrag ist ein Nachdruck aus dem Band »Das IW-Zukunftspanel: Ziele, Methoden, Themen und Ergebnisse«. Wir danken den Herausgebern für die freundliche Erlaubnis des Nachdrucks.

In diesem Kapitel werden die grundsätzlichen Probleme der Umfrageforschung – sogenannte Fehler – im Kontext des IW-Zukunftspanels angesprochen und Lösungsansätze dargestellt.

2. Das Fehlerkonzept in der Umfrageforschung

Wenn alle Objekte einer Grundgesamtheit befragt werden können (im Fall des IW-Zukunftspanels: alle Unternehmen der einbezogenen Branchen), sind die Verteilungen der interessierenden Merkmale und ihre Parameter bekannt – zum Beispiel vom Unternehmensumsatz der Durchschnitt oder Median über alle Unternehmen oder die Streuung als Maß der Unterschiedlichkeit der Umsätze. In aller Regel, so auch beim IW-Zukunftspanel, ist eine solche Vollerhebung jedoch nicht realisierbar, sondern es wird nur eine Teilmenge – eine Stichprobe – der Grundgesamtheit befragt.

Die Generalisierbarkeit der gewonnenen Daten auf die Grundgesamtheit (die inferenzielle Güte) sowie die Schlüssigkeit und die Validität der in einer Befragung erhobenen Angaben sind bei Teilerhebungen durch eine ganze Reihe von Fehlern gefährdet, auch Survey Errors genannt (Groves, 1989). In der Umfrageforschung werden vier potenzielle Fehlerquellen unterschieden:

- Coverage Error,
- Sampling Error,
- Nonresponse Error und
- Measurement Error.

Dieses Kapitel beschäftigt sich in erster Linie mit dem Nonresponse Error beim IW-Zukunftspanel. Die Fehlerarten werden im Folgenden aber kurz erläutert.

Coverage Error

Der Coverage Error ergibt sich daraus, dass eine Auswahlgesamtheit, die der Stichprobenziehung des Befragungsprojekts zugrunde liegt, zumindest teilweise nicht mit der angestrebten Zielpopulation (der Grundgesamtheit) übereinstimmt. Dadurch ist die inferenzielle Güte der Befragungsergebnisse gefährdet. Sind bestimmte Segmente einer Grundgesamtheit nicht oder in zu geringem Ausmaß in einer Auswahlgesamtheit enthalten, spricht man von Undercoverage. Im gegenteiligen Fall, der Überrepräsentation einer bestimmten Gruppe in der Auswahlgesamtheit, spricht man von Overcoverage (Schnell et al., 2008). Beim IW-Zukunftspanel kann davon ausgegangen werden, dass die Auswahlgesamtheit, das heißt alle in der auf Creditreform-Daten basierenden Unternehmensdatenbank MARKUS enthaltenen Unternehmen der Branchen Industrie (Verarbeitendes Gewerbe, Versorgung, Bau), Logistik und unternehmensnahe Dienstleistungen, ausreichend die Grundgesamtheit repräsentiert, also alle in Deutschland ansässigen Unternehmen dieser Branchen (vgl. Neligan & Schmitz, 2009).

Sampling Error

Wenn aus Stichprobendaten auf die unbekanntenen Werte der Grundgesamtheit geschlossen wird, spricht man in der Statistik von einer Schätzung. Eine solche Schätzung ist nur dann mathematisch handhabbar, wenn es sich bei der untersuchten Stichprobe um eine Zufallsauswahl handelt. Das heißt: Alle Elemente der Auswahlgesamtheit müssen eine bekannte (im einfachsten Fall: dieselbe) Wahrscheinlichkeit haben, in die Stichprobe zu gelangen. Die Konzepte der Schätztheorie, zum Beispiel die Erwartungstreuung und der Standardfehler, ermöglichen Aussagen darüber, wie gut eine Schätzung gelingt. Umgangssprachlich wird eine Zufallsstichprobe, die mit einer gewissen Güte die Schätzung von Parametern ermöglicht, als repräsentative Stichprobe bezeichnet. Für die Ergebnisse einer solchen Stichprobe kann also die Genauigkeit der Schätzung bestimmt werden. Die Ungenauigkeit dieser Schätzung wird als Sampling Error (Stichprobenfehler) bezeichnet. Er ist zum einen abhängig von der Größe der Stichprobe: Je größer die Zufallsstichprobe, desto geringer ist der Stichprobenfehler (Bortz, 1993). Zum anderen hängt er auch von der Technik der Stichprobenziehung ab (Bortz & Döring, 1995). Beim IW-Zukunftspanel handelt es sich um eine nach bestimmten Kriterien wie Umsatzgröße oder Branche geschichtete Stichprobe. Innerhalb dieser Schichten werden Unternehmen zufällig ausgewählt.

Nonresponse Error

Unabhängig davon, wie sorgsam die Auswahlgesamtheit definiert oder die Stichprobe gezogen wurde, antworten einige Mitglieder der Stichprobe nicht auf die Aufforderung, an einer Befragung teilzunehmen. Wenn sich diese sogenannten Nonresponder in den gemessenen Merkmalen von den Befragungsteilnehmern systematisch unterscheiden, wird dies zu einem Problem, das als Nonresponse Error oder Nonresponse-Fehler bezeichnet wird. Eine niedrige Ausschöpfungsquote – also ein niedriger Anteil der Befragten an der Zahl aller kontaktierten Unternehmen oder Personen – führt also nicht zwangsläufig zu einem hohen Nonresponse-Fehler. Dieser ist vielmehr abhängig von der Ausschöpfungsquote und der Unterschiedlichkeit in den gemessenen Merkmalen zwischen Responder und Nonresponder.

Die Ausschöpfungsquote stellt in der Regel – auch wenn das nicht unumstritten ist – implizit dennoch das einzige Maß für das Vorliegen eines Nonresponse-Fehlers dar. Deshalb ist man in der Markt- und Sozialforschung immer bestrebt, diese Quote zu maximieren, um dadurch Verzerrungen auf Grund von Nonresponse möglichst gering zu halten. Diese Form von Nonresponse wird auch als Unit-Nonresponse bezeichnet, da vollständige Units (beispielsweise Personen oder Unternehmen) ausfallen. Davon abgegrenzt wird das sogenannte Item-Nonresponse, das Fehlen von Angaben zu einzelnen Fragen. Eine feinere Unterteilung auf der Dimension von »keinerlei

Angaben« bis »vollständig vorliegende Angaben« im Kontext von Online-Befragungen findet sich bei Bosnjak (2002), ist aber an dieser Stelle nicht weiter relevant. Hier soll es allein um Betrachtungen zu Unit-Nonresponse gehen, weshalb der Begriff Nonresponse in diesem Kapitel synonym zu Unit-Nonresponse verwendet wird.

Measurement Error

Die drei bisher genannten Fehlerquellen resultieren alle daraus, dass bestimmte Elemente der Grundgesamtheit nicht befragt wurden. Daher werden sie auch als Errors of Nonobservation bezeichnet. Der Measurement Error ist dagegen dadurch definiert, dass Angaben der Befragten, also die beobachteten Werte, von den angenommenen wahren Werten der befragten Personen oder Unternehmen abweichen. Solche Fehler werden auch als Observational Errors bezeichnet. Dieses Konzept umfasst nicht nur unsystematische, zufällige Messfehler – diese allein würden sich über alle Befragten hinweg betrachtet beziehungsweise bei wiederholten Befragungen ausmitteln. Teil des Konzepts sind gerade auch systematische Antwortverzerrungen (sogenannte Biases). Groves (1989) nennt vier verschiedene Quellen des Measurement Errors:

- den Interviewer,
- das Instrument,
- den Befragten und
- die Befragungsmethode.

Fehler durch den Interviewer haben ihre Ursache zum Beispiel darin, dass Interviewer die Angaben fehlerhaft registrieren oder durch ihre Intonation beim Vorlesen den Befragten beeinflussen. Fehler aufgrund des Instruments rühren beispielsweise aus dem Layout des Fragebogens oder aus den verwendeten Formulierungen. Befragte wiederum können Fehlerquellen sein, weil sie sich zum Beispiel in ihrer Motivation oder in ihren kognitiven Fähigkeiten unterscheiden. Außerdem weisen verschiedene Befragungsmethoden bestimmte Charakteristika auf, welche die Angaben der Befragten beeinflussen können. Die vier Quellen von Measurement Errors können nicht als unabhängig voneinander gelten. So erscheint es plausibel, dass sich zum Beispiel die Einflüsse des Layouts eines Fragebogens oder auch bestimmter Charakteristika der Befragungsmethoden nicht unabhängig von den kognitiven Fähigkeiten oder der Motivation des Befragten auswirken. Es muss auch beachtet werden, dass Measurement Errors durch die Vielzahl an möglichen Ursachen besonders schwierig zu modellieren und zu messen sind und oft sehr restriktive Annahmen erfordern (Biemer & Stokes, 1991). Schließlich wurde der Measurement Error definiert als Abweichungen der gewonnenen Daten von wahren Werten. Die wahren Werte sind in der Regel nicht be-

kannt, schon gar nicht, wenn es sich um subjektive Phänomene wie Einstellungen, Erwartungen oder Ähnliches handelt, die erfasst werden sollen.

3. Ursachen von Nonresponse

Die Ursachen für Nonresponse lassen sich in Anlehnung an Bosnjak (2002) in drei Kategorien unterteilen:

- Nicht-Erreichbarkeit der Zielperson (beziehungsweise des Unternehmens),
- technische Einschränkungen und Störungen und
- sogenannte Noncompliance.

Unter Noncompliance versteht man sowohl die willentliche Teilnahmeverweigerung (entweder durch Ignorieren der Befragungseinladung oder durch aktive Verweigerung der Teilnahme) als auch die Unfähigkeit zur Teilnahme aus mangelnder Kompetenz.

Der Rekrutierungsprozess für das IW-Zukunftspanel findet in zwei Schritten statt (vgl. Neligan & Schmitz, 2009). Zunächst werden die Unternehmen telefonisch kontaktiert und um die Angabe einer E-Mail-Adresse und eines Ansprechpartners für die Befragung gebeten (Schritt 1). Anschließend wird an diese E-Mail-Adresse eine Einladung zur Teilnahme an der Befragung verschickt (Schritt 2). Sowohl beim telefonischen Erstkontakt als auch beim zweiten Kontakt per E-Mail kann Nonresponse aus den drei genannten Ursachen auftreten, wobei eine klare Trennung der Ursachen insbesondere beim Kontakt per E-Mail nicht möglich ist. Das soll am Beispiel folgender Befragungswellen exemplarisch erläutert werden:

- 4. Befragungswelle (Frühjahr 2007),
- 6. Befragungswelle (Winter 2007/2008) und
- 7. Befragungswelle (Frühjahr 2008).

Die Analyse des Teilnahmeverhaltens an diesem Rekrutierungsprozess beschränkt sich hier auf die in den jeweiligen Befragungswellen neu rekrutierten Teilnehmer. Nicht berücksichtigt werden Teilnehmer, die bereits an einer Vorwelle teilgenommen haben.

Schritt 1: Telefonischer Erstkontakt

Tabelle 2.1 gibt einen Überblick der erzielten Ergebnisse bei der telefonischen Kontaktierung in den Befragungswellen 4, 6 und 7:

- Immerhin 57,1 Prozent der angerufenen Unternehmen stellten ihre E-Mail-Adressen zur Verfügung, um eine Einladung zur Befragung zu erhalten.
- Die Nicht-Erreichbarkeit aufgrund falscher Telefonnummern und Kontakten ausschließlich zu Anrufbeantwortern lag durchschnittlich bei 14 Prozent der angerufenen Unternehmen.
- Technische Störungen spielten in dieser Phase praktisch keine Rolle.

- Immerhin durchschnittlich 25,5 Prozent der Unternehmen verweigerten jedoch an dieser Stelle die Teilnahme. Ebenfalls zu den nicht erreichbaren Unternehmen müssen solche gezählt werden, die angaben, über keine E-Mail-Adresse zu verfügen. Inwieweit diese Unternehmen eventuell doch eher den Verweigerern zuzurechnen sind, ist eine Frage, die noch offen ist und erst durch weitere Untersuchungen geklärt werden kann.

Tabelle 2.1: Ergebnisse des telefonischen Erstkontakts (Schritt 1) in Prozent der telefonisch kontaktierten Unternehmen

	4. Befragungs-welle	6. Befragungs-welle	7. Befragungs-welle	Durchschnitt
<i>E-Mail-Adresse erfasst</i>	61,4	55,9	54,4	57,1
<i>Abbruch durch technische Störung</i>	0	0	0	0
<i>Nicht erreichbar</i>	12,8	13,6	15,1	14,0
<i>Teilnahme verweigert</i>	22,7	24,6	27,9	25,5
<i>Keine E-Mail-Adresse</i>	3,0	5,9	2,6	3,3
<i>Insgesamt</i>	100	100	100	100
<i>Anzahl der Unternehmen</i>	12.378	6.033	16.939	35.350

Quelle: IW-Zukunftspanel 2007 und 2008, 4., 6. und 7. Befragungswelle

Schritt 2: Reaktionen auf E-Mail-Kontakt

In der zweiten Phase der Kontaktierung – per E-Mail – lag der Anteil der aktiven Verweigerer, die Rückmeldung darüber gaben, nicht an der Befragung teilnehmen zu wollen, bei durchschnittlich 8,8 Prozent (Tabelle 2.2). Dieser im Vergleich zum telefonischen Kontakt geringere Anteil liegt vermutlich zum einen daran, dass viele Unternehmen, die generell nicht an Befragungen teilnehmen, dies bereits beim telefonischen Kontakt mitteilten. Zum anderen muss auch bei den 55,9 Prozent, die gar nicht auf die Aufforderung zur Teilnahme reagierten, davon ausgegangen werden, dass ein erheblicher Anteil die Teilnahme schlicht durch Ignorieren der Einladung verweigerte. Einen vermutlich kleineren Anteil dieser 55,9 Prozent machen Unternehmen aus, welche die E-Mail übersahen oder deren Spamfilter die E-Mail gleich aussortierten. Schließlich finden sich darunter auch vereinzelt Unternehmen, die technische Schwierigkeiten beim Öffnen oder Bedienen des Fragebogens hatten, sich aber nicht bei der IW Consult oder dem Zentrum für Evaluation und Methoden der Universität Bonn meldeten. Meldeten sie sich doch, ließen sich die Störungen in der Regel schnell beseitigen. Es handelte sich dabei meist um Schwierigkeiten bei der Bedienung des Browsers oder um Einschränkungen des Internetzugriffs durch strenge Sicherheitsrichtlinien

in den Unternehmen. Nur in sehr wenigen Fällen traten Störungen durch die verwendete Befragungssoftware auf.

Es zeigt sich, dass eine klare Trennung der Ursachen für Nonresponse bei einem Kontakt per E-Mail schwierig ist. Anders sieht es aus bei eindeutig falschen E-Mail-Adressen, Fehlern bei der Übertragung der E-Mails oder Unzustellbarkeit aufgrund zu voller Postfächer, da der Absender hier eine automatische Rückmeldung erhält. Der Anteil unzustellbarer E-Mails lag beim ersten E-Mail-Kontakt durchschnittlich bei rund 10 Prozent. Zum Teil handelte es sich dabei um Erfassungsfehler der Interviewer, zum Teil aber auch um Fehler aufseiten der telefonisch Befragten. Durch Nachrecherchieren im Internet konnte dieser Anteil auf durchschnittlich 3,6 Prozent reduziert werden. Verglichen damit ist der Anteil von Abbrechern relativ hoch, vor allem in der 4. Befragungswelle mit 20,4 Prozent nicht abgeschlossenen Fragebögen. Dies lässt sich durch die Länge des Fragebogens in der 4. Welle erklären – der Fragebogen war vergleichsweise lang. Unternehmen benötigten zu seinem Abschluss im Durchschnitt ungefähr 31 Minuten, in den Befragungswellen 6 und 7 waren es demgegenüber nur 23 Minuten beziehungsweise 19 Minuten.

Die letztendliche Rücklaufquote, also der Anteil an Unternehmen, welche die Befragung komplett abschlossen, beträgt 14,7 Prozent. Diese auf den ersten Blick niedrige Quote ist für Online-Befragungen, die im Durchschnitt 11 Prozent niedrigere Rücklaufquoten aufweisen als andere Befragungsmethoden (Lozar Manfreda et al., 2008), noch nicht extrem niedrig, zumal hier auch unzustellbare E-Mails und wegen Dopplung entfernte Unternehmen berücksichtigt sind. Ohne diese liegt die Rücklaufquote bei 15,8 Prozent. Dieser Anteil erhöht sich noch einmal, wenn man die Teilnehmer berücksichtigt, die den Online-Fragebogen nicht abgeschlossen haben, aber hinreichend Angaben gemacht haben, sodass diese in der Auswertung berücksichtigt werden können (siehe Neligan & Schmitz, 2009, S. 21). Diese zählen hier als Teilmenge zu den Unternehmen, welche die Befragung begonnen haben.

Tabelle 2.2: Ergebnisse des E-Mail-Kontakts (Schritt 2) in Prozent der Unternehmen, deren E-Mail-Adresse telefonisch neu erfasst wurde

	4. Befragungswelle	6. Befragungswelle	7. Befragungswelle	Durchschnitt
<i>Teilnahme verweigert</i>	8,3	7,5	9,8	8,8
<i>Entfernt wegen Doppelung</i>	1,2	3,9	5,2	3,5
<i>E-Mail unzustellbar</i>	2,2	3,7	4,6	3,6
<i>Nicht reagiert</i>	54,1	63,5	54,7	55,9
<i>Befragung begonnen</i>	20,4	9,8	9,2	13,5
<i>Befragung abgeschlossen</i>	13,8	11,7	16,5	14,7
<i>Insgesamt</i>	100	100	100	100
<i>Anzahl der Unternehmen</i>	7.604	3.371	9.222	20.197

Quelle: IW-Zukunftspanel 2007 und 2008, 4., 6. und 7. Befragungswelle

Insbesondere Nonresponse aus Gründen von Noncompliance fällt ins Gewicht. Schließlich ist davon auszugehen, dass sich Personen und Unternehmen, die sich willentlich der Teilnahme an einer Befragung verweigern, am ehesten systematisch von solchen unterscheiden, die an der Befragung teilnehmen. Andere Fehler, beispielsweise technische Störungen oder auch Fehler bei der Erfassung von E-Mail-Adressen, treten dagegen unsystematisch auf. Der Nonresponse-Fehler resultiert – wie bereits oben erläutert – nicht notwendigerweise aus einer niedrigen Ausschöpfungsquote, sondern hängt von den Unterschieden zwischen Respondern und Nonrespondern ab.

Um Nonresponse-Fehler abschätzen zu können, ist es beispielsweise möglich, antwortende Unternehmen hinsichtlich ihrer Strukturmerkmale wie Branche oder Umsatzgröße mit Unternehmen zu vergleichen, welche die Teilnahme verweigerten oder auf die Aufforderung zur Teilnahme nicht reagiert haben. Darüber hinaus kann weiterer Aufwand betrieben werden, um die Nonresponder doch noch zu einer Teilnahme zu bewegen und deren Angaben mit den Antworten derjenigen Unternehmen zu vergleichen, die ohne so eine besondere Aufforderung teilnehmen. Eine solche Nonresponse-Untersuchung wurde im Rahmen des IW-Zukunftspanels im Sommer 2008 durchgeführt. Dabei wurden Unternehmen, welche die Teilnahme verweigert oder auf die Aufforderung zur Teilnahme gar nicht reagiert hatten, telefonisch noch einmal kontaktiert. Nachdem sie einige Kernfragen aus dem IW-Zukunftspanel beantwortet hatten, wurden die Ergebnisse mit denen aus der Online-Befragung verglichen, um den Nonresponse-Fehler schätzen zu können. Die Ergebnisse dieser Nonresponse-Untersuchung sind in Neligan & Schmitz (2009), S. 21f dargestellt.

4. Einzelne Korrelate und Determinanten von Noncompliance

In der Literatur finden sich – über die allgemeinen, rein deskriptiven Kategorien der Ursachen von Nonresponse hinaus – eine ganze Reihe von einzelnen Determinanten und Korrelaten, die mit dem Ausmaß an Noncompliance zusammenhängen. Bosnjak (2002) unterscheidet dabei zwischen befragungszentrierten und befragtenzentrierten Merkmalen, die an dieser Stelle ansatzweise dargestellt werden.

4.1 Befragungszentrierte Merkmale

Zu den befragungszentrierten Merkmalen zählen die Ankündigung der Befragung und die Implementierung des Befragungsprozesses. Auf einen weiteren Aspekt, die Gestaltung des Befragungsinstruments, wird hier nicht näher eingegangen.

4.1.1 Ankündigung der Befragung

Als Faktoren, welche die Teilnahmebereitschaft beeinflussen, haben sich der Auftraggeber (Sponsor) und der Verwendungszusammenhang einer Befragung erwiesen. Die dahinter stehende psychologische Variable ist nach Tourangeau et al. (2000) die wahrgenommene Legitimität und Wichtigkeit der Befragung, welche die Bereitschaft zur Teilnahme erhöht. Eine hohe Legitimität ist beispielsweise gegeben, wenn eine Befragung einen offensichtlichen gesellschaftlichen Nutzen hat oder durch eine öffentliche Einrichtung durchgeführt oder gefördert wird. Zum Beispiel erhielten Brunner & Carroll (1969) deutlich höhere Response-Raten, als sie eine Universität statt einer Marktforschungsfirma als Sponsor nannten. Heberlein & Baumgartner (1978) verglichen Response-Raten in 98 Studien und fanden höhere Raten bei von Behörden gesponserten Befragungen. Diese bei klassischen Papier- und Bleistift-Befragungen gefundenen Effekte sind auch für Online-Befragungen relevant (Bosnjak, 2002). Beim IW-Zukunftspanel wird die Legitimität der Befragung zum einen durch den Ruf des Instituts der Deutschen Wirtschaft Köln kommuniziert, zum anderen aber durch die explizite Nähe zur universitären Forschung in Form der Kooperation mit dem Zentrum für Evaluation und Methoden der Universität Bonn. Die Wichtigkeit der Befragung wird darüber hinaus noch durch den Absender der Befragungseinladung, Prof. Dr. Michael Hüther, Direktor des Instituts der Deutschen Wirtschaft Köln, vermittelt.

Ein weiteres Merkmal, das zumindest bei klassischen Papier- und Bleistift-Befragungen starken Einfluss auf die Teilnahmebereitschaft hat, ist die Personalisierung der Ansprache. Dazu gehören nach Dillman (2000) die namentliche Ansprache, ein Datum, ein namentlicher Ansprechpartner und eine persönliche – keine maschinelle – Unterschrift. Ob die persönliche Anspra-

che den gleichen Effekt auch bei Online-Befragungen zeitigt oder ob sie hier wegen Datenschutzbedenken sogar zum gegenteiligen Effekt führt, ist bisher nicht eindeutig geklärt. Erfahrungen im Kontext des IW-Zukunftspanels sprechen zumindest tendenziell für einen positiven Effekt: Beispielsweise haben in der 7. Befragungswelle von den personalisiert kontaktierten Unternehmen mit 18,8 Prozent deutlich mehr an der Befragung teilgenommen als von nicht personalisiert kontaktierten Unternehmen. Hier betrug der Anteil 8,7 Prozent. Allerdings sind diese Zahlen mit Vorsicht zu interpretieren, da es sich nicht um einen experimentellen Vergleich von personalisiert und nicht personalisiert kontaktierten Unternehmen handelt. So könnten Unternehmen, die tendenziell eher nicht an einer Befragung teilnehmen, unwilliger gewesen sein, einen persönlichen Ansprechpartner zu nennen.

Grundsätzlich wird im IW-Zukunftspanel beim telefonischen Kontakt versucht, einen persönlichen Ansprechpartner von den Unternehmen zu erhalten. In den E-Mails und auch im Fragebogen wird immer ein namentlicher Ansprechpartner bei der IW Consult aufgeführt. Die Angabe eines Datums erübrigt sich durch die automatische Datierung der E-Mail. Eine grafische Unterschrift dagegen empfiehlt sich nicht, da Bilder in E-Mails ein typisches Merkmal von Spam-E-Mails sind und ohnehin von vielen E-Mail-Programmen standardmäßig ausgeblendet werden.

Eine Vielzahl von Studien beschäftigt sich mit der Frage, wie sich die Auslobung verschiedener Belohnungen (Incentives) auf die Beteiligung an einer Umfrage auswirkt (zum Beispiel Porst, 1999; Singer et al., 2000). Im Kontext klassischer Papier-und-Bleistift-Befragungen zeigen Incentives in Form von vorab ohne Gegenleistung zur Verfügung gestellten kleineren Geldbeträgen die stärkste positive Wirkung auf die Teilnahmebereitschaft, was mit der sogenannten Reziprozitätsnorm (Cialdini, 1988) erklärt wird: Es entsteht bei den Befragten ein sozialer Druck zur Gegenleistung. Interessanterweise darf der Geldbetrag jedoch nicht zu hoch ausfallen, sonst ergibt sich ein gegenteiliger Effekt. Die Belohnung muss also in einem realistischen Verhältnis zu der erwarteten Gegenleistung (Ausfüllen des Fragebogens) stehen. Im Kontext von Web-Befragungen zeigten Bosnjak & Tuten (2001), dass monetäre Incentives zumindest dann nicht funktionieren, wenn diese nur in elektronischer Form (zum Beispiel via PayPal) zugestellt wurden. Ähnlich verhält es sich mit Gutscheinen. Das Geld muss auch in Web-Befragungen zum Anfassen sein, um als Incentive zu wirken (Birnholtz et al., 2004). Je nach Ausprägung des persönlichen Interesses bei der jeweiligen befragten Person kann schon allein die Rückmeldung der Untersuchungsergebnisse als – wenn auch weniger wirksames – Incentive wirken. Den Teilnehmern des IW-Zukunftspanels werden auf Wunsch sowohl allgemeine Ergebnisse der Befragung (zweiseitiger IW-Panel-Report) zur Verfügung gestellt als auch individuelle Auswertungen, in denen ihre eigenen Antworten den von ihnen selbst gewählten Vergleichsgruppen gegenübergestellt werden (IW-Panel-

Benchmarking) (siehe auch Abschnitt 1.8). Ob diese Ergebnisse und insbesondere der Vergleich mit den sogenannten Benchmarks als nichtmonetäre Incentives eine positive Wirkung auf die Teilnahmebereitschaft ausüben, muss Gegenstand weiterer Untersuchungen werden.

Als ein weiteres Merkmal, das die Teilnahmebereitschaft beeinflusst, gilt die angekündigte Dauer der Befragung. So berichteten Crawford et al. (2001) bei einer angekündigten Dauer von rund acht Minuten in einer Online-Befragung von einer um etwa 4 Prozent höheren Teilnahmebereitschaft gegenüber einer angekündigten Dauer von ungefähr 20 Minuten. Da die tatsächliche Dauer im Durchschnitt bei rund 20 Minuten lag, ist es jedoch nicht verwunderlich, dass in der Gruppe, die mit der eher unrealistischen Zeitangabe von nur acht Minuten geködert wurde, mehr Befragte vorzeitig abbrachen als in der anderen Gruppe.

Für das IW-Zukunftspanel ergab ein Experiment auf Basis einer frühen Befragungswelle ganz ähnliche Ergebnisse. In der 3. Befragungswelle im Sommer 2006 wurden dafür alle zu kontaktierenden Unternehmen zufällig in drei Gruppen unterteilt. In dieser Stichprobe waren sowohl neu rekrutierte Teilnehmer als auch Teilnehmer, die bereits an einer Vorwelle teilgenommen hatten. Die erste Gruppe erhielt keine Angabe über die Dauer der Befragung. Die zweite und die dritte Gruppe bekamen den Hinweis, dass die Befragung 25 beziehungsweise 35 Minuten dauere. Der Anteil derjenigen, die mit der Befragung begannen, lag bei der Gruppe ohne Zeitangabe signifikant am höchsten (36,7 Prozent), gefolgt von der Gruppe mit angekündigten 25 Minuten (31,1 Prozent) und der mit 35 Minuten (30,8 Prozent) (Abbildung 2.1). Die tatsächliche Befragungsdauer lag bei durchschnittlich 20,6 Minuten (Median: 16,2 Minuten). Die Anteile abgeschlossener Fragebögen fielen ebenfalls in der ersten Gruppe (ohne Zeitangabe) signifikant am höchsten aus (19,9 Prozent), gefolgt von Gruppe zwei (17,2 Prozent) und Gruppe drei (16,8 Prozent). Unterm Strich gab es damit keine Unterschiede in Bezug auf die Abbrecherquote.

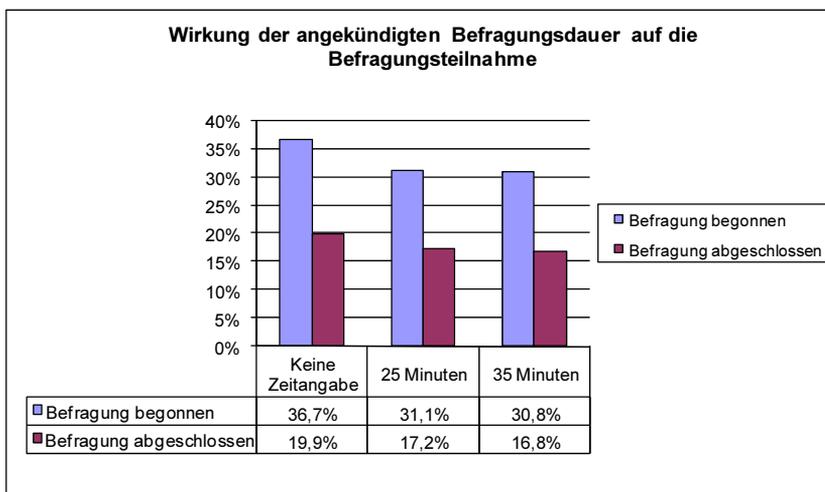


Abbildung 2.1: Wirkung der angekündigten Befragungsdauer auf die Befragungsteilnahme

Angaben in Prozent aller in Welle 3 angeschriebenen Unternehmen

Quelle: IW-Zukunftspanel 2006, 3. Welle

Das Fazit aus diesen beiden Untersuchungen zur angekündigten Dauer von Befragungen lautet: Steht die reine Maximierung der Ausschöpfungsquote im Vordergrund, sollte am besten auf eine Zeitangabe verzichtet werden. Aus Rücksicht auf die befragten Personen selber jedoch, die sich die Zeit zur Beantwortung des Fragebogens von ihrer Arbeitszeit oder Freizeit nehmen müssen, sollte eine ungefähre Angabe der zu erwartenden Befragungsdauer im Fragebogen enthalten sein. Sofern eine Angabe gemacht wird, sollte diese nicht unrealistisch niedrig, jedoch auch nicht höher als die durchschnittliche Befragungsdauer sein, insbesondere wenn es sich um eine längere Befragung handelt – wie beim IW-Zukunftspanel.

4.1.2 Erinnerungs-E-Mails

Die wirksamste Methode, die Ausschöpfungsquote bei postalischen Befragungen zu erhöhen, stellt zweifelsohne die wiederholte Kontaktierung der Zielpersonen mittels Erinnerungsschreiben (sogenannte Reminder) dar. Bei richtigem Vorgehen kann dies zu mehr als einer Verdopplung der Antwortraten führen (Dillman, 2000). Diese Wirkung der mehrmaligen Kontaktierung konnte aber auch für onlinebasierte Befragungen nachgewiesen werden (Cook et al., 2000).

Typischerweise werden neu rekrutierte Unternehmen während einer Befragungswelle bis zu viermal angeschrieben. Auf ein erstes formelles Anschreiben mit Absender Prof. Dr. Michael Hüther folgt wenige Tage später

eine etwas informellere E-Mail durch eine Mitarbeiterin, die im Auftrag von IW-Direktor Hüther nochmals um die Teilnahme an der Befragung bittet. Wenn auch nach etwa drei Wochen noch kein kompletter Fragebogen oder eine Verweigerung vorliegt, folgen das zweite und etwa zehn Tage später ein drittes und letztes Erinnerungsschreiben. In den verschiedenen E-Mails wird jeweils auch immer die Möglichkeit gegeben, durch Klick auf einen Link die Teilnahme an der Befragung zu verweigern. Diesen Unternehmen werden dann auch keinerlei weitere E-Mails zugesandt.

Bei Unternehmen, die schon mal an einer Befragung teilgenommen haben und ihre Bereitschaft zur Wiederteilnahme erklärt haben (Panelisten) wird ein ähnliches Vorgehen gewählt, nur dass auf das Anschreiben durch die Mitarbeiterin verzichtet wird. Somit werden Panelisten in einer Befragungswelle maximal dreimal zur Teilnahme aufgefordert.

Abbildung 2.2 zeigt die Wirkung der Erinnerungs-E-Mails auf die Rücklaufquote der neu rekrutierten Teilnehmer:

- Speziell der erste Reminder führt etwa zu einer Verdopplung der Rücklaufquote, bei den folgenden Remindern schwächt sich der Zuwachs leicht ab.
- Nach dem vierten Anschreiben haben schließlich rund 30 Prozent der neu rekrutierten Unternehmen mit der Befragung begonnen.

Zu beachten ist, dass hier auch Unternehmen gezählt werden, die den Fragebogen zunächst begonnen, dann aber die Teilnahme verweigert haben. Daher erscheint die Rücklaufquote hier etwas höher als zuvor in Tabelle 2.2 dargestellt.

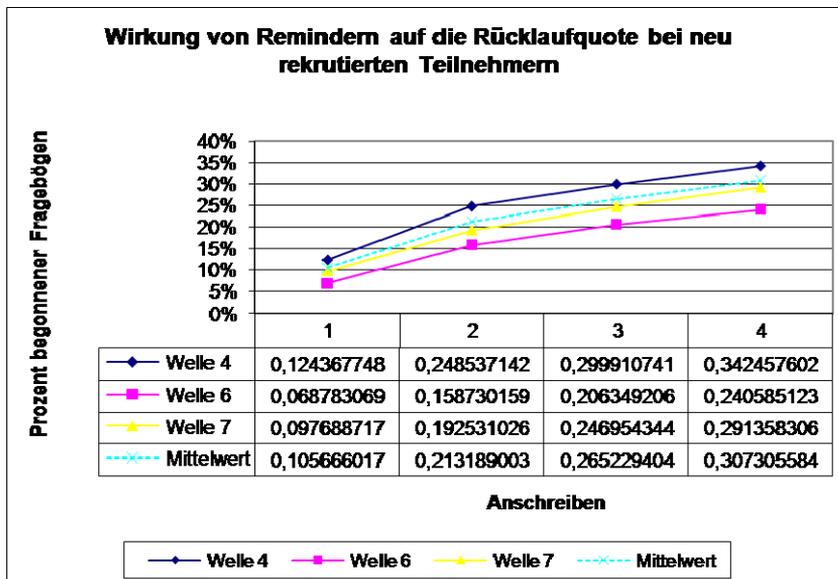


Abbildung 2.2 Rücklauf in Abhängigkeit der Anzahl der Anschriften Angaben der begonnenen Fragebögen in Prozent der telefonisch neu rekrutierten Unternehmen in Wellen 4, 6 und 7

Quelle: IW-Zukunftspanel, 2007-2008, 4., 6. und 7. Welle

4.2 Befragtenzentrierte Merkmale

Zu den befragtenzentrierten Merkmalen, die sich als Einflussfaktoren von Nonresponse erwiesen haben, zählen in Anlehnung an Bosnjak (2002):

- soziodemografische Merkmale der Befragten,
- Persönlichkeitsmerkmale der Befragten,
- das bisherige Teilnahmeverhalten,
- die Fähigkeiten und Fertigkeiten (Literacy), die zur Teilnahme an einer Befragung notwendig sind,
- die persönliche Relevanz beziehungsweise das Interesse für ein Befragungsthema,
- die (selbst berichteten) Teilnahmemotive sowie
- allgemeine, befragungsbezogene Einstellungen.

Obwohl es letztendlich Personen sind, die den Fragebogen des IW-Zukunftspanels ausfüllen, sind die relevanten Befragungseinheiten zunächst einmal Unternehmen. Daher wird an dieser Stelle auf Merkmale der Personen

– zum Beispiel Persönlichkeit, Teilnahmemotive oder Soziodemografie – nicht weiter eingegangen. Stattdessen werden die Merkmale der Unternehmen betrachtet. Eine Klassifikation nach Branche und Unternehmensgröße erscheint dabei sinnvoll. Diese Informationen liegen in Form der Einteilung in Wirtschaftszweige und Umsatzgrößenklassen für die Mehrzahl der Unternehmen (92 Prozent) vor. Tabelle 2.3 zeigt die Ausschöpfungsquoten in der ersten, telefonischen Rekrutierungsphase zusammengefasst als Durchschnitte für die Befragungswellen 4, 6 und 7.

Tabelle 2.3 Erfasste E-Mail-Adressen durch den telefonischen Erstkontakt (Schritt 1). Durchschnitt, in Prozent der telefonisch kontaktierten Unternehmen, mit Informationen über die Branchen- und die Umsatzgrößenklasse

	<i>Umsatz, in Millionen Euro</i>			
	<i>bis 1</i>	<i>über 1 bis 50</i>	<i>über 50</i>	<i>Insgesamt</i>
<i>Verarbeitendes Gewerbe</i>	47,9	65,7	70,9	57,3
<i>Logistik</i>	49,3	62,0	67,9	55,8
<i>Bau</i>	40,2	61,5	67,6	50,8
<i>Unternehmensnahe Dienstleistungen</i>	54,3	62,9	65,5	58,6
<i>Gesamt</i>	49,0	64,0	68,5	56,6

Verarbeitendes Gewerbe: inklusive Versorgung; Bau: Baugewerbe

Quelle: IW-Zukunftspanel 2007 und 2008, 4., 6. und 7. Befragungswelle

Im Durchschnitt der drei Befragungswellen gaben 56,6 Prozent der Unternehmen beim telefonischen Kontakt eine E-Mail-Adresse der Geschäftsführung an. Der Anteil der Unternehmen, die eine E-Mail-Adresse mitteilen, wächst dabei mit der Größe der Unternehmen. Bei einem Umsatz bis 1 Million Euro sind es nur 49 Prozent, bei den Unternehmen über 50 Millionen Euro Umsatz sind es 68,5 Prozent. Zudem unterscheidet sich die Ausschöpfungsquote auch zwischen den Branchengruppen. Am geringsten ist diese im Teilbereich Bau, speziell bei den kleineren Unternehmen. Dieser Zusammenhang findet sich dabei nicht nur in der zusammengefassten Auswertung, sondern auch in einer nach Befragungswellen getrennten Betrachtung.

Auch bei den Reaktionen auf die Einladung zur Befragungsteilnahme per E-Mail unterscheiden sich die Branchengruppen (Tabelle 2.4):

- Im Durchschnitt der drei Befragungswellen begannen 27,6 Prozent der Unternehmen, deren E-Mail erfasst wurde, mit der Bearbeitung des Fragebogens. In den Teilbranchen Bau und Verarbeitendes Gewerbe waren die Reaktionsquoten mit 19,7 beziehungsweise 25,1 Prozent unterdurchschnittlich hoch, im Bereich unternehmensnahe Dienstleistungen dagegen überdurchschnittlich (33,8 Prozent).

- Bezüglich der Unternehmensgröße zeigt sich in dieser Phase der Kontak-
tierung dagegen kein konsistenter Effekt auf die Teilnahmebereitschaft.

Tabelle 2.4 Begonnene Fragebögen als Reaktion auf E-Mail-Kontakt (Schritt 2)

Durchschnitt, in Prozent der Unternehmen, deren E-Mail-Adresse telefonisch neu erfasst wurde, mit Informationen über die Branchen- und die Umsatzgrößenklasse

	<i>Umsatz, in Millionen Euro</i>			
	<i>bis 1</i>	<i>über 1 bis 50</i>	<i>über 50</i>	<i>Insgesamt</i>
<i>Verarbeitendes Gewerbe</i>	27,4	23,3	24,9	25,1
<i>Logistik</i>	30,0	28,4	29,8	29,4
<i>Bau</i>	22,8	17,4	21,7	19,7
<i>Unternehmensnahe Dienstleistungen</i>	35,2	33,0	29,5	33,8
<i>Gesamt</i>	29,8	25,6	27,3	27,6

Verarbeitendes Gewerbe: inklusive Versorgung; Bau: Baugewerbe

Quelle: IW-Zukunftspanel 2007 und 2008, 4., 6. und 7. Befragungswelle

Das beschriebene Muster setzt sich zum Teil auch mit Blick auf die letztendlich abgeschlossenen Fragebögen fort (Tabelle 2.5):

- Auch hier schneiden die Teilbranchen Bau (10,5 Prozent) und Verarbeitendes Gewerbe (13,7 Prozent) unterdurchschnittlich ab, unternehmensnahe Dienstleistungen dagegen überdurchschnittlich (18 Prozent).
- Kleinere Unternehmen scheinen dagegen den Fragebogen eher bis zum Ende zu bearbeiten als größere.

Tabelle 2.5 Abgeschlossene Fragebögen als Reaktion auf E-Mail-Kontakt (Schritt 2)

Durchschnitt, in Prozent der Unternehmen, deren E-Mail-Adresse telefonisch neu erfasst wurde, mit Informationen über die Branchen- und die Umsatzgrößenklasse

	<i>Umsatz, in Millionen Euro</i>			
	<i>bis 1</i>	<i>über 1 bis 50</i>	<i>über 50</i>	<i>Insgesamt</i>
<i>Verarbeitendes Gewerbe</i>	15,6	12,6	11,6	13,7
<i>Logistik</i>	14,3	14,6	15,4	14,6
<i>Bau</i>	12,1	9,4	10,9	10,5
<i>Unternehmensnahe Dienstleistungen</i>	19,5	17,4	11,6	18,0
<i>Gesamt</i>	16,3	13,6	12,6	14,7

Verarbeitendes Gewerbe: inklusive Versorgung; Bau: Baugewerbe

Quelle: IW-Zukunftspanel 2007 und 2008, 4., 6. und 7. Befragungswelle

Es ist zu vermuten, dass sowohl die Bereitschaft, eine E-Mail-Adresse auf telefonische Anfrage herauszugeben, als auch die Bereitschaft, auf die Einladung zu einer Online-Umfrage zu reagieren, mit der Bedeutung und Verbreitung elektronischer Informationstechnologien im beruflichen Alltag zusammenhängen. Dies könnte die geringe Ausschöpfung bei kleineren Unternehmen der Baubranche erklären.

Da die hier dargestellten Unterschiede bereits in den ersten Panelwellen zu beobachten waren, wurden in den folgenden Befragungswellen Gegenmaßnahmen ergriffen. So gehen bei den Ziehungen der Telefonadressen (Bruttostichprobe) die verschiedenen Branchen und Unternehmensgrößen disproportional ein. Zudem werden für die Hochrechnung die Zellen repräsentativ gewichtet, um Nonresponse-Fehler bei den Ergebnissen des IW-Zukunftspanels zu minimieren.

5. Fazit

Die Aussagekraft von Befragungen im Allgemeinen – und damit auch des IW-Zukunftspanels im Speziellen – ist von einer Vielzahl an möglichen Fehlerquellen bedroht. Einige dieser Fehlerquellen (beispielsweise Interviewer-Effekte bei telefonischen Befragungen) sind seit Jahrzehnten in der Umfrageforschung bekannt, Maßnahmen zu ihrer Minimierung gehören zum etablierten Repertoire von Befragungsinstituten. Für Online-Befragungen als noch relativ neue Befragungsmethode gilt, dass insbesondere die Nonresponse-Problematik einer weiteren Betrachtung bedarf. Offen bleibt die Frage, ob bei einer Ausschöpfung von durchschnittlich knapp 15 Prozent (bei neu rekrutierten Unternehmen) nicht eine problematische Unterrepräsentierung einzel-

ner Segmente – zum Beispiel kleinerer Bauunternehmen – resultiert (vgl. Abschnitt 4.2), die langfristig nur unzureichend durch Überrepräsentation dieser Bereiche bei der Stichprobenziehung aufgefangen werden kann. Schließlich ist die Anzahl der Unternehmen in Deutschland für bestimmte, spezielle Branchen- und Umsatzgrößenklassen begrenzt.

In diesem Kapitel wurde eine ganze Reihe von Maßnahmen vorgestellt, mit denen versucht wird, die Ausschöpfungsquote zu erhöhen – zum Beispiel das mehrmalige Anschreiben, Incentives durch Bereitstellung von Befragungsergebnissen und die Betonung der Wichtigkeit der Befragung. Die Ergebnisse dieser Bemühungen lassen sich jedoch im Kontext onlinegestützter Unternehmensbefragungen wie dem IW-Zukunftspanel nicht mehr beliebig steigern. Aus diesem Grund kommt der Panelpflege, also der Erhöhung der Wiederbefragungsbereitschaft von bereits gewonnenen Panelteilnehmern, eine besondere Bedeutung zu. Die Frage, wie diese Bereitschaft langfristig gesichert werden kann, sollte Gegenstand zukünftiger Forschungsbemühungen sein.

Literatur

- BIEMER, P. & Stokes, L. (1991). Approaches to the modeling of measurement error: In Biemer, P., Groves, R.M., Lyberg, L., Mathiowetz, N. & Sudman, S (Hrsg.), *Measurement errors in surveys* (S. 487-516). New York.
- BIRNHOLTZ, J. P., Horn, D., Finholt, T., & Bae, S. J. (2004), The effects of cash, electronic, and paper gift certificates as respondent incentives for a web-based survey of technologically sophisticated respondents. *Social Science Computer Review*, 22 (3), 355-362.
- BORTZ, J. (1993). *Statistik für Sozialwissenschaftler*, Berlin: Springer
- BORTZ, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation für Sozialwissenschaftler*, Berlin: Springer.
- BOSNJAK, M. (2002). *(Non)Response bei Web-Befragungen*, Aachen: Shaker Verlag.
- BOSNJAK, M. & Tuten, T. L. (2001). Classifying response behaviours in web-based surveys, *Journal of Computer-Mediated Communication*, 6 (3). Zugriff am 24.11.2008 unter <http://jcmc.indiana.edu/vol6/issue3/boznjak.html>
- BRUNNER, G. A. & Carroll, S. (1969). The effect of prior notification on refusal rate in fixed address surveys, *Journal of Advertising*, 9 (1). 42-44.
- CIALDINI, R. B. (1988). *Influence: Science and practice*, Glenview (Illinois): Longman Higher Education.
- COOK, C., Heath, F. & Thompson, R.L. (2000). A meta-analysis of response rates in web- or internet based surveys, *Educational and Psychological Measurement*, 60, 821-836.
- CRAWFORD, S.D., Couper, M.P. & Lamias, M.J. (2001). Web surveys: Perception of burdens, *Social Science Computer Review*, 19 (2), 146-162.

- DILLMAN, D., (2000). *Mail and internet surveys: The tailored design method*, New York: John Wiley & Sons.
- GROVES, R. M., (1989). *Survey errors and survey costs*, New York: John Wiley & Sons
- HEBERLEIN, T. A. & Baumgartner, R. M., (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature, *American Sociological Review*, 43 (4),. 447-462.
- LOZAR Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I. & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates, *International Journal of Market Research*, 50 (1),. 79-104.
- NELIGAN, A. & Schmitz, E. (2009). Design und Analysepotenziale. In Lichtblau, K. & Neligan, A. (Hrsg.). *Das IW-Zukunftspanel – Ziele, Methoden, Themen und Ergebnisse* (S.11-33). Köln: Deutscher Instituts-Verlag.
- PORST, R. (1999). Thematik oder Incentives?: Zur Erhöhung der Rücklaufquoten bei postalischen Befragungen, In *ZUMA-Nachrichten*, 45. (23), 72-87.
- SCHNELL, R., Hill, P. & Esser, E. (2008). *Methoden der empirischen Sozialforschung*, München: Oldenbourg Verlag.
- SINGER, E., Hoewyk, J. van & Maher, M.P., (2000) Experiments with incentives in telephone surveys, *Public Opinion Quarterly*, 64 (4), 171-188
- TOURANGEAU, R., Rips, L. J. & Rasinski, K. (2000). *The psychology of survey response*, Cambridge University Press

Teil V

Perspektive

Das Zentrum für Evaluation und Methoden als Anbieter des Masterstudiengangs steht in diesem Teil im Vordergrund. Anschließend findet sich ein Überblick über *Umfrageprojekte* des ZEM. Zum Abschluss werden *Informationen über die Autor/inn/en* gegeben

Heike Steckhan, Katrin Wahlers & Kristina Heidemann

Wege zur Professionalisierung: Masterstudiengang Evaluation¹

Das Berufsfeld Evaluation besitzt aufgrund seiner Vielfalt einen »breit gefächerten Aufgabenkatalog« (Bortz & Döring, 2006, S. 97), der hohe und zum Teil konträre Anforderungen an Evaluatoren² und Evaluierete in sich birgt (Wottawa & Thierau, 2003), zumal der Anspruch an die Qualität von Evaluation steigt (Stockmann, 2004; Widmer 2004). Evaluation stellt in Deutschland bisher kein ausgewiesenes Berufsfeld dar (DeGEval, 2004), entsprechend sind auch die Ausbildungsmöglichkeiten eher rar und nicht systematisiert. In der Vergangenheit wurde daher bereits mehrfach angemahnt, dass die Situation des Tätigkeitsfeldes Evaluation eine Auseinandersetzung mit Fragen der Professionalisierung dringend notwendig mache, da dem steigenden Bedarf an Evaluation weder in quantitativer noch in qualitativer Hinsicht genügend personelle Ressourcen zur Verfügung stünden (Brandt, 2002; Heinrich & Scheffler, 2004; Schobert & Springer, 2006). Diesem Bedarf entsprechend hat die Universität Bern 2002 den ersten universitären Weiterbildungsstudiengang im deutschsprachigen Raum etabliert, innerhalb dessen auch einzelne Module am Studienstandort Köln angeboten werden. 2004 folgte ein Studienangebot in Kooperation der Universität des Saarlandes, der Hochschule für Technik und Wirtschaft des Saarlandes und der Katholischen Hochschule für Soziale Arbeit Saarbrücken. Ergänzend besteht seit dem Sommersemester 2008 auch in Bonn die Möglichkeit, den *Master of Evaluation* zu erwerben.

Wie aber soll ein Weiterbildungsangebot strukturiert sein, dass den Anforderungen der Teilnehmer, aber auch der Abnehmer gerecht wird? In der Literatur finden sich keine empirisch erhobenen Angaben zur Nachfrage nach Weiterbildung im Evaluationsbereich, weder von Seiten möglicher Zielgruppen noch von Expertenseite. Am Zentrum für Evaluation und Methoden (ZEM) wurden zwei Studien durchgeführt, welche dazu beitragen sollen, solche Fragen zu beantworten. Eine der Studien erfasste den bei Experten und potenziellen Studierenden vorhandenen Bedarf nach Weiterbildungsangeboten (Steckhan, 2006). Außerdem dienten die Antworten der Experten

1 Veröffentlicht in der Zeitschrift für Evaluation 2/2008 (S. 303-307)

2 Im Folgenden wird der besseren Lesbarkeit halber die geschlechtsneutrale oder männliche Formulierung verwendet, mit der aber Frauen und Männer gleichermaßen gemeint sind.

dazu, das Konzept einer Weiterbildung in Bezug auf Inhalt, Ablauf und Aufbau des Programms zu prüfen. Auf Basis dieses ersten Entwurfs wurden anschließend potenzielle Arbeit- und Auftraggeber von Evaluatoren zu ihrem Bedarf an professionell ausgebildeten Evaluatoren und ihren Anforderungen an die Ausbildung befragt (Wahlers, 2008). Ermittelt wurden so die Chancen der Absolventen auf dem Arbeitsmarkt bzw. die Branchen und Bereiche, in denen besonderer Bedarf an qualifizierten Evaluatoren besteht. Die Ergebnisse der beiden Arbeiten werden im Folgenden überblicksartig zusammengefasst.

Anforderungen an die Weiterbildung: Ein postgradualer Masterstudiengang Evaluation

Weiterbildung im Evaluationsbereich zielt vorwiegend auf eine Vermittlung von breit angelegten Qualifikationen ab, um Evaluatoren auf die unterschiedlichen Facetten des Tätigkeitsfeldes vorzubereiten. An die Weiterbildungsangebote wird daher der Anspruch gestellt, praxisnah zu sein und theoretische sowie methodische Qualifikationen zu vermitteln, die unmittelbar angewendet werden können (DeGEval, 2004).

Steckhan (2006) leitet aus den Ergebnissen ihrer Befragung ab, dass ein weiterbildender Masterstudiengang, der den Studierenden die Möglichkeit einer berufsbegleitenden Fortbildung gibt, ein geeignetes Instrument für die Qualifizierung von Evaluatoren darstellt: Die befragten Experten verlangten ein Weiterbildungsangebot, das sich an Akademiker verschiedener Fachbereiche richtet. Die in Bern, Saarbrücken und Bonn angebotenen Studiengänge entsprechen damit diesen grundlegenden Voraussetzungen. Entgegen der Empfehlung der befragten Experten wird in Bonn als Eingangsvoraussetzung für die Teilnahme am Masterstudiengang neben dem akademischen Abschluss einschlägige Berufserfahrung von mindestens einem Jahr gefordert. Dies lässt sich auf die hochschulrechtlichen Grundlagen für diese Form der Weiterbildung zurückführen.

Wahlers (2008) berichtet aus der Auseinandersetzung mit Arbeit- und Auftraggebern, dass häufig eine unzureichende methodische Ausbildung sowie mangelnde Sozial- und Selbstkompetenzen bei den in der Praxis agierenden Evaluatoren zu finden seien. In den Pflichtmodulen eines Masterstudiengangs soll deshalb ein möglichst umfassender Überblick über die Grundlagen von Evaluation (Theorie, Methodik, Umgang mit erhobenen Daten etc.) vermittelt werden, außerdem praktische Kompetenzen wie z.B. Budgetkalkulationen und -verhandlungen, Umgang mit Kritik und Formulierung von Ansprüchen an Evaluationen und ihre Konsequenzen eingeübt werden. Für Auftraggeber interessant ist die Bewertung von Evaluationsangeboten. Die Wahlpflichtmodule bieten die Möglichkeit zur Vertiefung von Wissen in spezifischen Themenfeldern. Wie von den Experten gefordert, besteht

die Gruppe der Dozenten aus Personen mit sowohl praktischem als auch theoretischem Bezug zum Thema Evaluation.

Grundsätzlich verfügt ein Masterstudiengang über den gewünschten Anwendungsbezug: Neben der konsequenten Erläuterung der theoretisch vermittelten Inhalte anhand praktischer Beispiele, beinhalten die Masterstudiengänge im deutschsprachigen Raum auch ein Praxismodul, das den Studierenden die Möglichkeit bietet, im eigenen Umfeld aber auch in anderen Bereichen das Gelernte umzusetzen. Darüber hinaus bietet die Masterarbeit die Möglichkeit zur Auseinandersetzung mit praktischen Inhalten.

Bedarf an qualifiziert ausgebildeten Evaluatoren

Die von Steckhan (2006) befragten Evaluationsexperten schätzen die Nachfrage nach einer formalisierten Weiterbildung als hoch ein. Aufgrund der mangelnden Qualifizierung und Professionalisierung im Evaluationssektor wurde die Ausbreitung entsprechender Studiengänge als notwendig bezeichnet. Die Ergebnisse der ebenfalls im Rahmen dieser Studie durchgeführten Studierendenbefragung bestätigen dies: 40% der befragten Bonner Studierenden gaben an, Interesse an Weiterbildungsangeboten zum Thema Evaluation zu haben, 25% beurteilten konkret die Form eines postgradualen, berufsbegleitenden Masterstudiengangs positiv.

Wahlers (2008) macht deutlich, dass auch seitens der Auftrag- und Arbeitgeber ein Bedarf an qualifiziert ausgebildeten Evaluatoren besteht, wobei die Hälfte der Befragten diesen Bedarf an Evaluatoren auch für die eigene Organisation bestätigt. Vor allem die Mitglieder der Branchen *Aus-, Weiterbildung und Unterricht* sowie *Forschung und Wissenschaft* sehen diesen Bedarf. Er wird von Seiten des öffentlichen Diensts und von Non-Profit-Organisationen eher bestätigt als von der Privatwirtschaft. Für die Zukunft prognostiziert die Mehrheit der befragten Auftrag- und Arbeitgeber einen Anstieg des Bedarfs an professionell ausgebildeten Evaluatoren. In den Branchen *Aus-, Weiterbildung und Unterricht, Forschung und Wissenschaft* sowie *öffentliche Verwaltung* trifft dies in der Mehrheit der Fälle auch auf die eigene Organisation zu.

Nutzen eines Masterabschlusses Evaluation für die potentiellen Teilnehmer

Ob ein Masterabschluss Evaluation berufliche Vorteile wie verbesserte Aufstiegs- und Einstiegschancen eröffnet, lässt sich noch nicht eindeutig beantworten, da die Transparenz der durch die Studiengänge in Saarbrücken, Bern und Bonn vermittelten Qualifikationen für die Arbeit- und Auftraggeber im Bereich Evaluation aufgrund des relativ kurzen Bestehens solcher Angebote noch nicht ausreicht. Die Ergebnisse von Wahlers (2008) legen aber

nahe, dass ein solcher Abschluss tendenziell berufliche Vorteile zu eröffnen scheint.

Insgesamt stimmten jeweils etwa ein Drittel der in der Studie von Wahlers (2008) befragten Arbeit- und Auftraggeber der Aussage zu, dass Evaluatoren mit einem entsprechenden Masterabschluss bei ansonsten gleichen Voraussetzungen bei der Auftragsvergabe bevorzugt würden, eher in einer Organisation eingestellt würden sowie erhöhte Chancen auf beruflichen Aufstieg hätten. Es wurde betont, dass nicht unbedingt ein formaler Abschluss, sondern konkret die Qualifikationen einer Person für derartige Entscheidungen relevant sei. Der Rückschluss von einem formalen Abschluss auf die Qualifikationen einer Person scheint demnach für die Arbeit- und Auftraggeber im Bereich Evaluation nicht zwingend gegeben zu sein. Im Hinblick auf diese Fragestellung bleibt vermutlich abzuwarten, bis sich Absolventen eines Masterstudiengangs Evaluation in der beruflichen Praxis bewährt haben und dadurch deren Qualifikationen für Arbeit- und Auftraggeber im Bereich Evaluation deutlich werden.

Unterstützungsbereitschaft der Organisationen für Mitarbeiter, die sich für eine Teilnahme an einem Weiterbildungsstudiengang Evaluation entscheiden, lässt sich bei einer Vielzahl der von Wahlers (2008) befragten Arbeitgeber feststellen. Teilnehmer eines solchen Studiengangs können hierbei insbesondere auf flexible Handhabung ihrer Arbeitszeiten zählen und somit die berufsbegleitende Weiterbildungsmaßnahme in Einklang mit ihrer Tätigkeit absolvieren. Zur Übernahme von Weiterbildungskosten oder Freistellung von der beruflichen Tätigkeit ist ebenfalls ein Großteil der Befragten bereit.

Fazit

Als Ergebnis beider Studien wurde deutlich, dass ein weiterbildender Masterstudiengang Evaluation bei Experten, potenziellen Studierenden sowie Arbeit- und Auftraggebern begrüßt wird, was insgesamt gute Chancen für solche Weiterbildungsangebote offenbart. Deutlich wurde, dass vor allem Sozial- und Selbstkompetenzen sowie umfassende und praktisch bedeutsame Methodenkenntnisse vermittelt werden müssen. Darüber hinaus ergaben sich Hinweise darauf, dass der öffentliche Dienst sowie Non-Profit-Organisationen vor allem aus den Branchen *Aus-, Weiterbildung und Unterricht, Forschung und Wissenschaft* sowie *öffentliche Verwaltung* einen allgemein steigenden Bedarf an qualifizierten Evaluatoren, insbesondere aber auch für ihre eigene Organisation sehen, so dass auch die Berufsaussichten der Absolventen positiv eingeschätzt werden können. Durch die Anpassung des Studiengangs in Bonn an die von Experten, Studierenden und Empfängern geäußerten Anforderungen erfolgt die Ausbildung möglichst nah an den Bedürfnissen des Marktes und vereinfacht es den Teilnehmern, nach ihrem Abschluss eine Anstellung zu finden.

Literatur

- BORTZ, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4., überarbeitete Auflage). Heidelberg: Springer.
- BRANDT, T. (2002). *Qualifikationsanforderungen für Evaluatoren. Überlegungen zur Entwicklung eines Ergänzungsstudiums Evaluation*. Unveröffentlichte Diplomarbeit, Freie Universität Berlin.
- DEUTSCHE Gesellschaft für Evaluation. (2004). *Empfehlungen für die Aus- und Weiterbildung in der Evaluation – Anforderungsprofile an Evaluatoreninnen und Evaluatoren*. Zugriff am 28.02.2008 unter <http://www.degeval.de/calimero/tools/proxy.php?id=51>
- HEINRICH, T. & Scheffler, D. (2004). Aus- und Weiterbildung in der Evaluation – Aktivitäten des Arbeitskreises der Deutschen Gesellschaft für Evaluation. *Zeitschrift für Evaluation*, 2, 351-355.
- SCHOBERT, B. & Springer, E. (2006). Erfassung von Aus- und Weiterbildungsangeboten und Informationen über Praktikamöglichkeiten mit deutlichem Evaluationsbezug durch die DeGEval. *Zeitschrift für Evaluation*, 2, 365-368.
- STECKHAN, H. (2006). *Eine Ex-Ante-Evaluation eines wissenschaftlichen Weiterbildungsangebots Evaluation*. Unveröffentlichte Diplomarbeit, Rheinische Friedrich-Wilhelms-Universität Bonn.
- STOCKMANN, R. (2004). Evaluation in Deutschland. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 13-43). Opladen: Leske + Budrich.
- WAHLERS, K. (2008). *Erwartungen an einen postgradualen Studiengang im Bereich Evaluation – eine Ex-Ante-Evaluation*. Unveröffentlichte Diplomarbeit, Rheinische Friedrich-Wilhelms-Universität Bonn.
- WIDMER, T. (2004). Qualität in der Evaluation – Wenn Wissenschaft zur praktischen Kunst wird. In R. Stockmann (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (S. 83-109). Opladen: Leske + Budrich.
- WOTTAWA, H. & Thierau, H. (2003). *Lehrbuch Evaluation*. Bern: Hans Huber.

Umfrageprojekte des ZEM von A bis Z¹

Accenture

Entwicklung, Durchführung und Auswertung einer Befragung unter den Media-Verantwortlichen der Top 300 Werbetreibenden in Deutschland zu unterschiedlichen Vermarktungsmodellen.

http://www2.wuv.de/w_v_infocenter/studien/vermarktungsstudie_werbungtreibende_nehmen_intransparenz_in_kauf

Telefonische und online Erhebung unter Projektmanagern zum Thema Erfolgsfaktoren von Konvergenzprodukten.

Telefonische und Online-Erhebung in deutschen und ausgewählten internationalen Unternehmen zur Einbindung von softwareintegrierten Produkten in ihren Unternehmen.

http://www.gespraechskreis.unipassau.de/fileadmin/Dateiendownload/Accenture_Studie_2009_Softwareintegrierte_Produkte_in_der_Fertigungswirtschaft.pdf

Entwicklung, Durchführung und Auswertung von Haushaltsbefragungen zu aktuellen Fragestellungen der Medienforschung, z.B. repräsentative Befragung von Internetnutzern zum Verbreitungsgrad von Videoportalen und deren Nutzungsformen.

http://www.presseportal.de/pm/39565/1077516/accenture_gmbh

Alanus Hochschule

Implementierung von Modul- und Lehrveranstaltungsevaluation sowie allgemeiner Studierendenbefragung gemäß dem »Bonner Modell der Hochschulevaluation« an der Alanus Hochschule.

Alexander von Humboldt-Stiftung

Interne Evaluation des »Transatlantic Science and Humanities Program« (TSHP) mittels einer Onlinebefragung sowie halbstandardisierter Interviews.

1 Für diverse Auftraggeber (Auswahl in alphabetischer Reihenfolge) werden ausgewählte Projekte genannt.

AQAS e.V. (Agentur für Qualitätssicherung durch Akkreditierung von Studiengängen)

Halbjährliche Onlinebefragung der Akkreditierungsgutachter und Ansprechpartner an Hochschulen zwecks Evaluation des Akkreditierungsverfahrens.

Auswärtiges Amt in Zusammenarbeit mit der Robert Bosch Stiftung und der Deutschen Gesellschaft für Auswärtige Politik (DGAP)

Externe Evaluation des Diplomatenkollegs mittels Alumnibefragung (online), Gruppendiskussion und persönlicher Interviews.

Auswärtiges Amt

Evaluation und Qualitätssicherung des *Besucherreiseprogramms* des Auswärtigen Amtes.

b1 communication

Onlinebefragungen von Architekten zu den Themen Leuchten, Sanitärmarken und Dämmung.

Online-Befragung von Elektroinstallateuren zum Thema Leuchten.

Beierlein Marketing / Habitat

Kundenzufriedenheitsbefragung (face to face) in fünf deutschen Habitat-Märkten.

Bundesanstalt für Straßenwesen (BASt)

Wiederholte Befragung von Fahranfängern zu ihren Meinungen und Einstellungen zu Themen der Verkehrssicherheit.

Durchführung und Auswertung von qualitativen Interviews zur Praktikabilität und Alltagstauglichkeit alkoholsensitiver Zündsperrern (Alcolocks) im Lastkraftverkehr zu drei Messzeitpunkten.

Bundesinstitut für Berufliche Bildung (BiBB)

Entwicklung einer Machbarkeitsstudie für die Etablierung eines Kommunikations- und Informationssystems für die berufliche Bildung (KIBB) sowie Durchführung von Interviews und Expertenworkshops und Entwicklung von Szenarien.

Auswertung und Aufbereitung von Daten zu verschiedenen Fachschulfächern an Berufsschulen.

CEWS

Balancierung von Wissenschaft und Elternschaft (BAWIE): Recherche von ca. 40.000 Email-Adressen von Mitarbeitern an 19 deutschen Universitäten, Durchführung einer Onlinebefragung zum Thema Elternschaft und Auswertung der Daten u.a. mittels AnswerTree und Clusteranalyse.

Unterstützung der internen Evaluation diverser Förderprogramme von Frauen in Wissenschaft und Forschung.

DAAD

Anpassung der bestehenden Hochschulevaluation der Universität Bonn an die besonderen Bedürfnisse ausländischer Studierender.

DEKRA

Reevaluation eines Kursprogramms zur Wiederherstellung der Kraftfahrreignung.

Deutscher Entwicklungsdienst (DED)

Methodische Betrachtung der Auswahlstrategien für Fachkräfte der Entwicklungshilfe.

Deutscher Hochschulverband

Methodische Beratung, Umsetzung und Auswertung einer schriftlichen Befragung zur Erfassung der aktuellen Fundraising Situation an deutschen Hochschulen und Medizinischen Fakultäten.

Entwicklung eines Rankings für Hochschulrektoren und dessen erstmalige Umsetzung im Rahmen einer Online-Befragung.

Ernst-Moritz-Arndt-Gymnasium Bonn

Onlinebefragung von Schülern, Lehrern, Eltern und Ehemaligen des Gymnasiums analog zum Bonner Modell der Hochschulevaluation.

eye square

Telefonische Akquise von Teilnehmern für verschiedene Laborstudien und Umsetzung verschiedener Onlinebefragungen mit bis zu 10.000 Teilnehmern im Bereich Marktforschung.

Fraunhofer Institut für angewandte Informationstechnik

Durchführung und Auswertung einer Befragung von IT-Verantwortlichen kleiner und mittelständischer Unternehmen zum Thema *Usability*, 300 Interviews.

Telefonische Haushaltsbefragung zur Nutzung von Handys und E-Mails.

goetzpartners

Durchführung und Auswertung einer telefonischen Haushaltsbefragung zum Thema »Kenntnis und Interesse an den technischen Möglichkeiten des Fernsehens der Zukunft«. 2008 als Hybridstudie (telefonische Akquise von E-Mailadressen für die eigentliche Online-Studie), 2009 reine CATI-Studie.

Heel

Quartalsweise Erhebung der Verkaufszahlen ausgewählter Produkte der am Heel-Partnerprogramm teilnehmenden Apotheken (Panel mit n=1.500 Apotheken), online und telefonisch.

Hertie Stiftung Stipendiatenwerk

Erstellung eines Resümees zum zehnjährigen Bestehen des Roman-Herzog - Stipendiatenprogramms für Nachwuchswissenschaftler aus Mittel- und Südosteuropa.

IBM Business Consulting

Entwicklung, Umsetzung und Auswertung einer Online-Befragung von über 1.500 Praktikern weltweit aus unterschiedlichen Branchen zum Thema »Change Management«: Making Change Work. Continuing the enterprise of the future conversation.

<http://www-935.ibm.com/services/us/gbs/bus/pdf/gbe03100-usen-03-making-change-work.pdf>

http://www-935.ibm.com/services/de/bcs/pdf/2009/making_change_work.pdf

http://www-05.ibm.com/de/pressroom/downloads/mcw_2007.pdf

Entwicklung, Durchführung und Auswertung einer telefonischen Haushaltsbefragung zur Erwartungshaltung von Privatpersonen gegenüber öffentlicher Verwaltung und Sozialversicherungen (n=1.000).

Entwicklung, Durchführung und Auswertung einer telefonischen Unternehmensbefragung von Einkäufern und Anbietern von »Wissensintensiven Dienstleistungen«.

Entwicklung, Durchführung und Auswertung von Haushaltsbefragungen zu aktuellen Fragestellungen der Medienforschung, z.B. Veränderungen der Mediennutzungsmuster junger Rezipienten bei zunehmender Konvergenz der Medien.

<http://www-05.ibm.com/de/pressroom/downloads/medienstudie-2008.pdf>

<http://www-05.ibm.com/de/pressroom/downloads/energie-studie.pdf>

[http://www.hdm-](http://www.hdm-stutt-)
[stutt-](http://www.hdm-stutt-)

[gart.de/wi/is/forschung_und_transfer/broadcast_infotage/broadcast_infotag08/scherf_rolle_konsumenten.pdf](http://www.hdm-stutt-gart.de/wi/is/forschung_und_transfer/broadcast_infotage/broadcast_infotag08/scherf_rolle_konsumenten.pdf)

Entwicklung, Durchführung und Auswertung einer telefonischen Befragung von norddeutschen Unternehmen zum Thema Energieeffizienz und Klimaschutz.

iFQ

Entwicklung, Durchführung und Auswertung einer Befragung in Zusammenhang mit dem Monitoring der Exzellenzinitiative des Bundes und der Länder.
http://www.forschungsinfo.de/Publikationen/Download/working_paper_5_2008.pdf

Impuls GmbH

Durchführung und Auswertung telefonischer Haushaltsbefragungen zu den Themen Alkoholkonsum, Geld und Glück, Genuss und Verantwortung.
http://www.bonne-nouvelle.de/downloads/geld_macht_gluecklich.pdf
http://www.bonne-nouvelle.de/downloads/gesundheit_und_lebensqualitaet.pdf

Durchführung und Auswertung einer telefonischen Unternehmensbefragung zum Thema Corporate Social Responsibility (CSR).
<http://www.bonne-nouvelle.de/downloads/csr-studie-2009.pdf>

Umsetzung diverser Mitarbeiterbefragungen bzw. Führungskräftebefragungen.

Durchführung und Auswertung einer telefonischen Haushaltsbefragung zum Thema Konsummuster, Konsumsituationen, Konsummotive und erlebte Wirkung von Alkohol (Alkoholkonsum und Alkoholsozialisation).

Evaluation eines Raucherentwöhnungsprogramms (Happy End) mittels einer Teilnehmerbefragung zu fünf Messzeitpunkten.

InWent

Durchführung einer Tracer-Study mit dem Ziel, den beruflichen Verbleib ehemaliger International Leadership Training (ILT)-Programmteilnehmer zu dokumentieren und den aktuellen Weiterbildungsbedarf der Alumni zu erfassen.

Evaluation des Arbeits- und Studienaufenthalt-Programms (ASA) und seiner Umsetzung.

Alumnibefragungen zu den Stipendienprogrammen Asia/Pazifik und USA; Onlinebefragungen.

Entwicklung, Durchführung und Auswertung einer Online-Befragung brasilianischer Alumni von InWent-Trainingsmaßnahmen.

IW Köln

Aufbau eines Unternehmenspanels und vierteljährliche Online-Befragung von Geschäftsführern zu aktuellen Fragen des Strukturwandels: IW Köln Zukunftspanel.

Telefonische Akquise von Emailadressen und Online-Befragung ostdeutscher Maschinenbau-Unternehmen zur Zukunft Ihres Wirtschaftszweiges.
Mitarbeiterbefragungen (Papier und online) verschiedener Unternehmen.
Telefonische Unternehmensbefragung zum Thema Vereinbarkeit von Familie und Beruf (n=1.150).
Telefonische Unternehmensbefragung zum Thema Local Sourcing und Einkauf.
Telefonische Verbandsbefragung zur Bewertung des Einsatzes eines InfoMobils für die Metall- und Elektrobranche.
Telefonische Befragung von Unternehmen, die in der Vereinigung Bergischer Unternehmensverbände organisiert sind.
Koordination einer telefonischen Unternehmensbefragung zur Erfassung von Angebot und Nutzung familienfreundlicher Maßnahmen in sechs europäischen Ländern: Familienmonitor 2009.
Durchführung und Auswertung einer telefonischen Befragung von Unternehmen in deutschen Großstädten zur Reformfreudigkeit ihrer Stadt. Jährlich wiederholte Erhebung mit etwa 3.000 Interviews.
<http://www.wiwo.de/politik/leben-arbeit-zukunft-50-deutsche-staedte-im-vergleich-305964/>

King's College in London, Department of Palliative Care, Policy & Rehabilitation

Durchführung einer Onlinebefragung im Rahmen des EU-Projektes PRISMA unter Palliativmedizinern zu Erfahrungen mit Messinstrumenten und Fragebögen.

Kreispolizeibehörde Rhein-Sieg-Kreis

Projektelevaluation: Ermittlung von Beweggründen für Fehlverhalten von Radfahrern sowie von geeigneten Maßnahmen zur Bewirkung von Verhaltensveränderungen im Radverkehr im Rahmen eines Projektes zur Erhöhung der Verkehrsdisziplin von Fahrradfahrern.

Lilly

Decision Support System for Hospital Pharmacies (DSSHP): Befragung von Krankenhaus-Apotheken zu Organisation und Ablauf von Bestellungen.

LVR

Repräsentative Besucherbefragung aller Museen des LVR.

Mars Drinks Germany

Erhebung der Kundenzufriedenheit (Unternehmen) mit KLIX - Getränkeautomaten.

media.lysis GmbH

Vergleich der Schätzgenauigkeit längs- und querschnittlicher Erhebungen sowie Untersuchung zur Güte telefonischer, erinnerungsbasierter Hörverhaltensforschung.

Orchesterzentrum NRW

Evaluation der Lehre und des Studienangebots des Orchesterzentrum NRW.

Robert Bosch Stiftung

Externe Evaluation des Stipendienprogramms zur Förderung von amerikanischem Führungsnachwuchs mittels Alumnibefragung (online), Auswertung von Sekundärdaten und persönlichen Interviews.

Robert Bosch Stiftung & Studienstiftung des deutschen Volkes & Auswärtiges Amt

Externe Evaluation des Stipendienprogramms des Stiftungskollegs für internationale Aufgaben mittels Alumnibefragung (online), Auswertung von Sekundärdaten und persönlichen Interviews.

Stadt Bonn, Wirtschaftsförderung

Befragung junger, in der Ausbildung befindlicher Menschen verschiedener Bildungszweige mit dem Ziel der Einschätzung ihrer Eignung für unternehmerische Selbstständigkeit sowie einer Gegenüberstellung von Selbstbild und Fremdbild der Unternehmerpersönlichkeit.

Stifterverband für die deutsche Wissenschaft

Schriftliche Befragung von Hochschulen, Forschungseinrichtungen, Verbänden und Unternehmen zum Thema Innovationsfaktor Kooperation, d. h. Kooperationsbeziehungen zwischen Wissenschaft und Wirtschaft. Über eine Internetplattform wurden den Teilnehmern ein Papierfragebogen, eine Onlinebefragung sowie die Möglichkeit zum Up- und Download eigener Dokumente zur Verfügung gestellt. Vorbild der Studie war das »Lambert Review of Business«.

Onlinebefragung von Hochschulen, Stiftungsprofessoren, Unternehmen, Stiftungen und Ministerien zum Thema Stiftungsprofessuren an deutschen Hochschulen. http://www.stifterverband.org/publikationen_und_podcasts/positionen_dokumentation/stiftungsprofessuren_in_deutschland/stiftungsprofessuren_in_deutschland.pdf

Durch eine schriftliche Befragung und ergänzende qualitative Interviews mit zentralen Ansprechpartnern in Unternehmen wurden die aktuellen Anteile ausländischer Akademiker, insbesondere der Anteil ausländischer Forscher am FuE-

Personal, Rekrutierungsstrategien sowie Gründe und Motive der Unternehmen erfasst, die zu einem Einsatz ausländischer Akademiker in Forschung und Entwicklung führen können.

SUZ Sozialwiss. Umfragezentrum Duisburg

Befragung von Privathaushalten in 26 hessischen Gemeinden zur Ermittlung der Einstellung zur EU und der Determinanten dieser Einstellung.

T-Com

Benchmarkanalyse des Kundenservice der acht größten deutschen Kommunikationsanbietern per Mystery-Tests der Kommunikationskanäle Internet, Telefon und Face-to-Face (im Shop).

Technische Universität München (TUM): Agentur Mädchen in Wissenschaft und Technik

Verbleibstudie der Teilnehmerinnen des Programms »Mädchen machen Technik«.

T-Mobile International

TMO ICCA Study: Gender-related satisfaction at Contact Centres & Shops: Statistische Reanalyse von Studien zur Kundenzufriedenheit in Abhängigkeit vom Geschlecht der Kunden und des TMO Service aus sechs europäischen Ländern.

Koordination und Auswertung einer Benchmark-Studie »Kundenzufriedenheit« in zehn europäischen Ländern.

Brand Strength: Benchmark-Studie zum Thema Markenbekanntheit sowie Markenstärke von Mobilfunkanbietern. Repräsentative telefonische Befragung von Mobilfunknutzern in neun europäischen Ländern.

Telefonische Haushaltsbefragung zur Markenbekanntheit und Markenstärke von Mobilfunkanbietern, insgesamt 72.000 Interviews.

trinloc

Bevölkerungsrepräsentative Studie zur Bekanntheit und Nutzung von Digitalradio (DAB).

<http://www.dr->

[m.info/Meinungsbarometer/download/Meinungsbarometer_Maerz_2007.pdf](http://www.dr-m.info/Meinungsbarometer/download/Meinungsbarometer_Maerz_2007.pdf)

TÜV Nord Mobilität GmbH & Co. KG

Evaluation von SPEED-02 (Sicherheit durch Prävention: Erfahrungen mit und Engagement gegen Drogen)

Universität Bonn

BISS (Bonn International Scholar Service): Evaluation des Angebots des Welcome Centres, einer Beratungs- und Servicestelle für international mobile Forscherinnen und Forscher an der Universität Bonn.

Onlinebefragung zur Betreuungssituation ausländischer Studenten an der Universität Bonn.

Online-Erfassung der universitären Angebote im Rahmen der »Jungen Uni«

Onlinebefragung im Rahmen der Erstsemesterbefragung 2009 zu Bekanntheit und Nutzung des Studienkompass.

Universität Bonn & Wirtschaftsförderung der Stadt Bonn

Entwicklung, Durchführung und Auswertung einer Onlinebefragung von Mitarbeitern und Spinoffs der Universität Bonn sowie einer persönlichen Befragung von Dekanen mit dem Zweck der Bedarfsanalyse eines Gründer- und Innovationszentrums.

Universität Bonn, Geographisches Institut

Evaluation des von der EU geförderten Programms post-graduate training school mit Einbezug der Studierenden und Dozenten.

Universität Bonn, Gleichstellungsbüro

Evaluation des Mentoring- & Trainingsprogramms für Nachwuchswissenschaftlerinnen der Universität Bonn.

Universität Bonn, Institut für Landtechnik, Haushaltstechnik

Erhebung bei deutschen Privathaushalten zur möglichen Einführung eines Frischeindikators auf leicht verderblichen Lebensmitteln. Der Indikator macht eine Unterbrechung der Kühlkette für den Verbraucher sichtbar.

<http://www.pr-inside.com/de/bonner-studie-ermittelt-anwendun-r1210712.htm>

Universität Bonn, Institut für Wissenschaft und Ethik

Entwicklung, Durchführung und Auswertung einer telefonischen Haushaltsbefragung zum Umgang mit Bluttests am Beispiel der Alzheimer Demenz, zudem Erfassung psychometrischer Tests.

Entwicklung, Durchführung und Auswertung einer telefonischen Haushaltsbefragung zur Möglichkeit der deutlichen Verlängerung der menschlichen Lebensspanne und der Verlangsamung des Alterungsprozesses.

Universität Bonn, Pressestelle

Begleitung des Projektes über den gesamten Projektverlauf, regelmäßige Befragungen der Portalnutzer (online) und Journalisten.

Universität Bonn, Rektorat

Konzeption, Durchführung und Auswertung einer Onlinebefragung zu Mobilitätsverhalten und Mobilitätsbedürfnissen unter allen Mitarbeitern und Studierenden der Universität Bonn sowie telefonischer Interviews mit Ansprechpartnern zentraler Institute. Die Ergebnisse der Befragungen dienen der Konzeption eines Mobilitätsplans für die Universität Bonn, welcher eine Grundlage für die zukünftigen Bauprojekte der Universität darstellt.

Universität Hannover, Institut für Wirtschafts- und Kulturgeographie & Institut für Arbeitsmarkt- und Berufsforschung (IAB)

Durchführung einer telefonischen Haushaltsbefragung zu Gründungsaktivitäten im Rahmen des internationalen Global Entrepreneurship Monitor (GEM)-Projektes (n=6.000).
http://www.gemconsortium.org/files.aspx?Ca_ID=114

Universität Leipzig

An der Universität Leipzig führt das ZEM in einem wissenschaftlichen Kooperationsprojekt die Hochschulevaluation nach dem »Bonner Modell der Hochschulevaluation« durch. Dabei kommen alle Module des Bonner Modells zum Einsatz: Modul- und Lehrveranstaltungsevaluation, Allgemeine Studierendenbefragung, Absolventenbefragung sowie die Befragung des wissenschaftlichen Personals.

Verband der Vereine Creditreform e.V.

Durchführung und Auswertung von Telefonbefragungen zur Bewertung der Finanzmarktkrise durch den Mittelstand; Konjunkturmfrage unter Versandhändlern (Mitglieder des BVH Bundesverband des Deutschen Versandhandels, jährlich).

BWHT Konjunkturmfrage (Panel): Durchführung und Auswertung einer quartalsweise stattfindenden Konjunkturbefragung in Baden-Württemberg. Zielgruppe: verschiedene Handwerksgruppen

Durchführung und Auswertung zweier telefonischer Umfragen (Anbieter- & Nachfragerseite) zum Thema »industrielle Dienstleistungen«.

VolkswagenStiftung

Es werden Szenarien der Mobilitätskultur im Jahr 2030 entwickelt. Dabei stehen die »zukünftigen« älteren (d. h. in 2030: 65+) Menschen im Mittelpunkt. Die Beteiligung der (zukünftigen) Älteren wird hier u. a. durch systematische Berücksichtigung ihrer Zukunftsentwürfe und heutigen Lebensgewohnheiten und Lebensstile im Rahmen einer repräsentativen Befragung der »Babyboomer« Generation gewährleistet (n=1000).

Wissenschaftsrat

Technische Umsetzung der Wissenschaftsratings (1. Welle: Chemie, Soziologie; 2. Welle: Geschichte, Elektrotechnik)

Zentrum für Entwicklungsforschung ZEF

Konzeption einer Teilnehmerbefragung in zwei vom ZEF durchgeführten Trainingskursen innerhalb des Projekts WISDOM (Wasserinformations-System zur nachhaltigen Entwicklung des Mekong Deltas in Vietnam).

Zeitungs Marketing Gesellschaft (ZMG)

Durchführung und Auswertung einer telefonischen Befragung unter Neuabonnenten einer lokalen Tageszeitung.

Zentrum für europäische Wirtschaftsforschung (ZEW)

Telefonische Unternehmensbefragung im Rahmen der Gründerinitiative 'unternehm was'. Erfassung von Trends und Entwicklungsperspektiven bei Hightech-Unternehmen (n=1.000).

<ftp://ftp.zew.de/pub/zew-docs/gutachten/hightechgruendungen.pdf>

Evaluation des Förderprogramms »EXIST 3« des Bundesministeriums für Wirtschaft und Technologie (BMWi): Telefonische Befragung von jungen Unternehmen zur Bedeutung von Hochschulen und öffentlichen Forschungseinrichtungen für Unternehmensgründungen.

Aufbau eines Panels neu gegründeter Unternehmen. Jährliche Befragung zur wirtschaftlichen Situation, zur Auftragslage und Unternehmensentwicklung, sowie zur Unternehmensfinanzierung.

<http://www.zew.de/de/forschung/projekte.php3?action=detail&nr=816>

Telefonische Befragung von forschungsintensiven Unternehmen, die im Rahmen des BMBF- Förderprogramms »KMU-innovativ« eine Projektskizze eingereicht haben mit dem Ziel der Evaluation des Programms.

Telefonische Befragung von geschlossenen Unternehmen zur Ursache für das Scheitern von jungen Unternehmen in den ersten fünf Jahren ihres Bestehens. N etwa 4.000.

Durchführung und Auswertung einer telefonischen Befragung unter jungen Hightech-Unternehmen, die im Rahmen der Befragung »unternimm was« im Jahr 2007 angesprochen wurden. Thema: Internationalisierung/ Vertriebsstrategien junger Unternehmen auf ausländischen Märkten.

Zentrum für Management- und Personalberatung

Umsetzung einer Onlinebefragung von Mitarbeitern und Führungskräften im Rahmen eines 180 Grad Feedback-Verfahrens zur Führungsstil-Analyse.

Autorenverzeichnis

- BORSCH, Alexandra: Diplom-Psychologin. Von 2004 bis 2007 im Telefonlabor des ZEM tätig; Diplomarbeit zum Net Promoter Score. Seit 2007 Personalberaterin bei der SCG SchielConsulting GmbH. E-Mail: alexandra.borsch@gmx.de
- COSTA, René: Diplom-Psychologe & Magister Artium. Seit Anfang 2008 als Wissenschaftlicher Mitarbeiter am ZEM im Bereich Hochschulevaluation beschäftigt. E-Mail: costa@zem.uni-bonn.de
- DONAGHY, Dr. Kieran P.: Professor an der Cornell University, Ithaca, NY, USA. Head of the Dept. of City and Regional Planning. Forschungen u. a. in den Bereichen Mobilität, Verkehr und Nachhaltigkeit, Modellierung regionaler und globaler polit-ökonomischer Entscheidungen.
- FEESER-Lichterfeld, Dr. Uwe: Diplom-Theologe & Diplom-Psychologe. Seit 2005 Wissenschaftlicher Mitarbeiter im Institut für Wissenschaft und Ethik an der Universität Bonn, wissenschaftliche Bearbeitung und Koordination des von der NRW-Landesregierung geförderten Projekts »Lebensverlängerung und Verlangsamung des menschlichen Alterns.«. E-Mail: feeser@iwe.uni-bonn.de
- FUCHS, Dr. Michael: Geschäftsführer des Instituts für Wissenschaft und Ethik an der Universität Bonn. Leiter der Studie »Lebensverlängerung und Verlangsamung des menschlichen Alterns« zusammen mit Prof. Dr. Georg Rudinger. E-Mail: fuchs@iwe.uni-bonn.de
- HEIDEMANN, Dr. Kristina: Diplom-Psychologin. Seit 2001 am ZEM tätig. Projektleiterin mit den Schwerpunkten Hochschulevaluation, Evaluation von Bildungs- und Lehrzielen, Mobilitäts- und Verkehrserziehung, Auslandsaufenthalte und Koordinatorin des Masterstudiengangs Evaluation an der Universität Bonn. E-Mail: heidemann@zem.uni-bonn.de
- HILGER, Dr. des. Norbert: Diplom-Psychologe. Seit 2001 am ZEM in den Bereichen Umfragen und Statistik beschäftigt. Derzeitige Tätigkeitsfelder sind Konzeption und Analyse von Umfragen und Evaluationsstudien, Betreuung von Diplom- und Masterarbeiten, Lehre in den Bereichen Statistik und Evaluation am Institut für Psychologie und im Masterstudiengang Evaluation der Universität Bonn. E-Mail: hilger@zem.uni-bonn.de
- HÖRSCH, Katharina: Diplom-Psychologin. Seit 2005 am ZEM tätig. Projektleiterin mit den Schwerpunkten Hochschulevaluation, Schulevaluation und Evaluation von Akkreditierungsverfahren, Organisation der ZEM-Fachtagungen. E-Mail: hoersch@zem.uni-bonn.de
- HUBER, Stefan: Diplom-Psychologe Wissenschaftlicher Mitarbeiter an der FU Berlin im DFG Projekt: »Gesundheitskompetenz: Modellentwicklung und Validierung«. E-Mail: stefan.huber@fu-berlin.de

- KLEINEMAS, Dr. Uwe: Diplom-Psychologe. Mitarbeiter in der Abteilung Methodenlehre & Diagnostik des Instituts für Psychologie. Seit Februar 2002 Geschäftsführer des Zentrums für Alternskulturen (ZAK). Langjährige Kooperation mit dem ZEM durch Lehrveranstaltungen für angehende Diplompsychologen. E-Mail: u.kleinemas@uni-bonn.de
- KOCHERSCHIED, Dr. Kristina: Diplom-Psychologin. Seit 2003 am ZEM und seit 2009 am Zentrum für Alternskulturen (ZAK) tätig. Projektleiterin mit den Schwerpunkten Verkehr und Mobilität im Alter, Aus- und Weiterbildung für Evaluatoren, Evaluation und qualitative Methoden. E-Mail: kocherscheid@zak.uni-bonn.de
- KRÜGER, Dr. Thomas: Diplom-Psychologe Seit 2002 Leiter des Bereichs »Umfragen & Statistik« des ZEM. Arbeitsschwerpunkte sind Umfrageforschung, Evaluation und statistische Modellierung. E-Mail: krueger@zem.uni-bonn.de
- LENARTZ, Norbert: Diplom-Psychologe. Seit November 2007 Wissenschaftlicher Mitarbeiter an der FU Berlin im Projekt »Gesundheitskompetenz: Modellentwicklung und Validierung«. E-Mail: norbert.lenartz@fu-berlin.de
- MAYERBÖCK, Astrid: Diplom-Psychologin. Seit 2003 am ZEM tätig. Projektleiterin im Bereich Internationale Marktforschung mit den Schwerpunkten Kundenzufriedenheit und Kundenbindung. E-Mail: mayerboeck@zem.uni-bonn.de
- PAULUS, Michael: Geschäftsführer Ericsson Services GmbH. E-Mail: michael.paulus@ericsson.com
- PIETRANGELI, Sandra: Diplom-Psychologin. Seit Anfang 2008 am ZEM tätig. Projektleiterin mit den Schwerpunkten Change Management, Kundenzufriedenheit und Kundenbindung sowie Online Self-Assessment. E-Mail: pietrangel@zem.uni-bonn.de
- PONERT, Maximilian: Diplom-Psychologe Seit 2009 am ZEM als Projektleiter im Bereich Evaluation tätig. E-Mail: ponert@zem.uni-bonn.de
- POPPELREUTER, Dr. Stefan: Diplom-Psychologe. Bereichsleiter für die Impuls GmbH. In zahlreichen Projekten, u.a. zu den Themen Nachhaltige Mobilität und Verkehrssicherheit von SeniorInnen, Zusammenarbeit mit dem ZEM. E-Mail: Poppelreuter@impuls-gmbh.com
- REITER, Dr. Anton: Studium der Physik und Promotion (in Kernphysik) an der Universität Frankfurt am Main, Tätigkeiten in verschiedenen naturwissenschaftlichen Fachverlagen, von 1991 bis 2006 Geschäftsführer der TÜV Media GmbH.
- RIEK, Simon: Diplom-Psychologe. Seit 2002 am ZEM tätig. Als IT- und Methoden-Spezialist leitet er den Bereich »Online-Erhebungen« und ist für die Sicherheit und Funktion der IT-Infrastruktur des ZEM zuständig. Sein Forschungsschwerpunkt liegt auf Online-Befragungsmethodik. E-Mail: riek@zem.uni-bonn.de
- RUDINGER, Prof. Dr. Georg: Professor für das Fach Psychologie am Institut für Psychologie der Universität Bonn: Abt. Methodenlehre, Diagnostik und Evaluation., Forschungsschwerpunkte: Statistische Modelle, Methoden und Evaluation; Technik, Neue Medien, Mobilität (und Alter); Entwicklungspsychologie

über die Lebensspanne. Prof. Rudinger ist Leiter des ZEM und des Zentrums für Alternskulturen (ZAK, gegr. 2002). Seit 2007 ist Prof. Rudinger Rektorsbeauftragter für Qualitätssicherung und Evaluation an der Universität Bonn. E-Mail: rudinger@uni-bonn.de

SINDERN, Eva: Diplom-Psychologin. Seit 2006 am ZEM tätig. Projektleiterin mit den Schwerpunkten Verkehrs- und Mobilitätserziehung, Kulturpolitik und Online Self-Assessment. E-Mail: sindern@zem.uni-bonn.de

SOELLNER, Prof. Dr. Renate: Seit Juni 2003 Juniorprofessorin für Evaluation, Qualitätssicherung und Qualitätsmanagement in Erziehungswissenschaft und Psychologie an der FU Berlin, ab WS 2009/10 an der Universität Hildesheim. Innerhalb des DFG-Schwerpunktprogramms »Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen« Leitung des Projektes »Gesundheitskompetenz: Modellentwicklung und Validierung« mit Prof. Dr. Rudinger. E-Mail: soellner@zedat.fu-berlin.de

STECKHAN, Heike: Diplom-Psychologin. Seit 2002 am ZEM tätig. Projektleiterin in den Bereichen Bildung, Kultur- und Kulturpolitik und Koordinatorin des Masterstudiengangs Evaluation an der Universität Bonn.
E-Mail: steckhan@zem.uni-bonn.de

WAHLERS, Katrin: Diplom-Psychologin. Verfasste 2007 ihre Diplomarbeit zum Thema Masterstudiengang Evaluation.