

Matthias D. Mahlendorf

Eskalation des Commitments bei scheiternden Projekten

Matthias D. Mahlendorf

Eskalation des Commitments bei scheiternden Projekten

Eine empirische Untersuchung kognitiver Eskalationsfaktoren und
verhaltenswissenschaftlich basierter Controllingmaßnahmen

Mit einem Geleitwort von Prof. Dr. Dr. h. c. Jürgen Weber

Dissertation, Wissenschaftliche Hochschule für Unternehmensführung (WHU) – Otto Beisheim
School of Management; Vallendar 2008

Eskalation des Commitments bei scheiternden Projekten

Alle Rechte vorbehalten

© 2008 Matthias D. Mahlendorf

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechts ist ohne vorherige schriftliche Zustimmung des Autors unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Druck: Hamacher, Vallendar

Printed in Germany

ISBN 3-937141-37-5

Geleitwort

Das Thema der vorliegenden Dissertation ist innerhalb der Forschung am Institut für Management und Controlling an der WHU (IMC) in das Feld der Rationalitätssicherung einzuordnen. Mit der Verlusteskalation wird ein in der Praxis beobachtbares Rationalitätsdefizit angesprochen, das darin besteht, trotz besserer möglicher Einsicht zu lange an nicht (mehr) erfolgversprechenden Projekten festzuhalten. Die Gründe für ein solches Verhalten sind ganz offensichtlich nicht im Modell des homo oeconomicus zu finden, sondern erfordern eine verhaltenswissenschaftliche Öffnung der Analyse. Für die empirische Relevanz der Verlusteskalation liegt eine Vielzahl von Studien und praktischen Erfahrungen vor. Für einen Forschungszeitraum, der sich mit Formen und Ursachen von Rationalitätsdefiziten und Maßnahmen für ihre Vermeidung, Verminderung oder Beseitigung beschäftigt, ist Verlusteskalation ein besonders reizvolles Feld, weil die für dieses Phänomen anfälligen Projekte i. d. R. durch hohe Unsicherheit gekennzeichnet sind. Insofern ist die Grenze zwischen einem zu langen Festhalten an Projekten und einem „zu frühen ins Korn Werfen der Flinte“ eher vage.

Das Thema Verlusteskalation ist in der Forschung am Institut nicht neu, sondern bereits in der Dissertation von *Zayer* ausführlich konzeptionell bearbeitet worden. Der Forschungsstand ist zudem durch eine intensive empirische Forschung gekennzeichnet, die allerdings bislang fast ausschließlich experimenteller Natur war. Bei dieser ergibt sich gerade wegen der schwierigen Nachbildbarkeit der Kontextbedingungen für eine Verlusteskalationssituation die sehr berechtigte Frage nach der Inhaltsvalidität der Experimente. Großzahlige empirische Forschung, die das Phänomen breiter erfassen kann und Respondenten befragt, die mit dem Phänomen in der täglichen Praxis tatsächlich Erfahrung haben (und nicht – wie Laborexperimente – mangels entsprechendem fehlenden Zugang zu Praktikern auf Studenten zurückgreift), fehlte bisher.

Exakt hier setzt die Arbeit von *Mahlendorf* an. Sie verwendet die konzeptionellen Überlegungen und Ergebnisse von *Zayer* als Basis und baut darauf eine großzahlige empirische Erhebung auf, die mit einer Zahl von 677 Respondenten und einer Rücklaufquote von 11,6 Prozent eine zugleich repräsentative wie umfangreiche Datengrundlage besitzt. Die Erhebung ist darüber hinaus durch eine geradezu mustergültige methodische Sorgfalt gekennzeichnet. *Mahlendorf* berücksichtigt erstmals in Studien des Instituts – und meines Kenntnisstandes nach gilt dies für empirische Controllingstudien im deutschsprachigen Raum generell – das Phänomen der sozialen Erwünschtheit von Antworten explizit, indem er hierzu eine eigene Skala zur Überprüfung entwickelt. Zudem erfolgt eine Kontrolle durch die parallele Befragung von Managern und Controllern. Auch der mögliche Einfluss von Freundschaft wird kontrolliert. Dieses Vorgehen hebt

sich in Sorgfalt und Tiefe deutlich vom bisherigen Standard empirischer Studien im Controlling ab.

Im Ergebnis zeigen die empirischen Ergebnisse im Vergleich zu den bisher vorliegenden experimentellen Erkenntnissen zum einen Übereinstimmung – nämlich in der Wirkung der untersuchten Einflussfaktoren auf den Umfang der Projekteskalation –, zum anderen aber deutliche Abweichungen: Insbesondere hinsichtlich der Gegenmaßnahmen werden die Experimente nur zu einem geringen Teil bestätigt. Hier wird sowohl ein erheblicher weiterer Forschungsbedarf sichtbar als auch deutlich, wie wichtig es ist, unterschiedliche empirische Forschungsmethoden miteinander zu verbinden, um der Realität gerecht zu werden. Nicht nur deshalb ist der Arbeit von *Mahlendorf* ein großer Leserkreis zu wünschen.

Prof. Dr. Dr. h. c. Jürgen Weber

So I have just one wish for you – the good luck to be somewhere where you are free to maintain the kind of integrity I have described and where you do not feel forced by a need to maintain your position in the organization or financial support, or so on, to lose your integrity. May you have that freedom.

Richard P. Feynman

Danksagung

Mein Dank gehört all jenen, die mir in den vergangenen Jahren die von Richard Feynman beschriebene Freiheit ermöglicht haben. Dabei ist an erster Stelle mein Doktorvater Prof. Dr. Dr. h. c. Jürgen Weber zu nennen, der die Arbeit nicht nur fachlich betreut hat, sondern auch ein hervorragendes Forschungs- und Arbeitsumfeld geschaffen hat und gleichzeitig in vielerlei Hinsicht persönliche Vorbildwirkung ausübte. Bedanken möchte ich mich auch bei Prof. Dr. Andreas Hoffjan für die Übernahme des Zweitgutachtens und das hilfreiche und motivierende Feedback. Dr. Carl Marcus Wallenburg gilt mein großer Dank für vielfältige forschungsmethodische Anregungen, seine konstruktive Kritik sowie die, trotz hoher Leistungsanforderungen, stets lockere und humorvolle Arbeitsatmosphäre.

Ganz im Sinne des Zitats von Anselm Feuerbach „Einen Menschen den man lieb hat und eine große Idee, die die Seele ausfüllt, was braucht man weiter?“ hatte ich das große Glück, neben der Dissertation – als die Seele ausfüllende Idee – gleich eine ganze Reihe von lieben Menschen um mich zu haben: zuallererst natürlich meine geliebte Frau; darüber hinaus aber auch meine Eltern, Großeltern, Schwiegereltern und deren ganze Familie, Freunde aus der Schul- und Studienzeit sowie meine Großtante, die 50 Jahre vor mir ebenfalls kognitive Aspekte erforscht hat (vgl. Rieden 1959). Bei ihnen allen möchte ich mich für die erfahrene Zuneigung und Unterstützung von ganzem Herzen bedanken. Ein großes Dankeschön richtet sich auch an alle Kollegen; es freut mich sehr, dass sich hier Freundschaften entwickelt haben, die hoffentlich weit über die Promotionszeit hinaus Bestand haben werden.

Die vorliegende Schrift wurde Ende 2008 an der WHU – Otto Beisheim School of Management eingereicht und als Dissertation angenommen. Die Erstellung der Arbeit wurde durch ein Promotionsstipendium der Friedrich-Naumann-Stiftung mit Mitteln des Bundesministeriums für Bildung und Forschung gefördert.

Matthias D. Mahlendorf

Inhaltsübersicht

1	Einleitung	1
2	Theoretische Basis der Untersuchung	9
3	Methodik und Durchführung der empirischen Studie	45
4	Statistische Analysen und Diskussion der Ergebnisse	95
5	Zusammenfassung und Fazit	201
	Anhang	213
	Literaturverzeichnis	241

Inhaltsverzeichnis

Inhaltsübersicht	VIII
Inhaltsverzeichnis	XI
Tabellenverzeichnis	XIX
Abbildungsverzeichnis	XXIII
Abkürzungsverzeichnis	XXV
1 Einleitung	1
1.1 Das Phänomen der Eskalation des Commitments	1
1.2 Relevanz der Arbeit	2
1.2.1 Relevanz für die Wissenschaft	2
1.2.2 Relevanz für die Praxis	5
1.3 Zielsetzung der Arbeit	7
1.4 Aufbau der Arbeit	7
2 Theoretische Basis der Untersuchung	9
2.1 Konzeptionelle Grundlagen	9
2.1.1 Entwicklung der Behavioral-Accounting-Forschung und Einordnung der Arbeit	9
2.1.2 Begriffsklärungen und Fokus der Arbeit	12
2.1.3 Verknüpfung der psychologischen Erklärungen mit der ökonomi- schen Theorie	15

2.1.4	Abgrenzung von nahe stehenden Arbeiten	18
2.2	Herleitung des Modells	21
2.2.1	Das Modell von Zayer als Ausgangspunkt	23
2.2.2	Modifikation des Modells für die empirische Prüfung	23
2.2.2.1	Geltungsbereich des Modells	23
2.2.2.2	Modifikation des Entscheidungsprozessmodells	25
2.2.2.3	Reduzierung der Elemente	26
2.2.3	Das theoretische Modell für die empirische Prüfung	27
2.3	Erläuterung der Variablen und Herleitung der Hypothesen	29
2.3.1	Eskalationstendenz und Projekteskalation	29
2.3.2	Wahrnehmungsdefizite	30
2.3.2.1	Wahrnehmungsschwelle	30
2.3.2.2	Selektive Wahrnehmung	34
2.3.3	Änderung des Zielsystems	35
2.3.3.1	Self-Justification-Effekt	35
2.3.4	Bewertungsfehler	38
2.3.4.1	Sunk-Cost-Effekt	38
2.3.4.2	Optimismus	40
2.3.5	Moderierende Faktoren	41
2.3.5.1	Zeitpunkt der Wahrnehmung	41
2.3.5.2	Bereitschaft, Freiraum auszunutzen	42
2.3.5.3	Overconfidence	43
3	Methodik und Durchführung der empirischen Studie	45
3.1	Diskussion der Datenerhebung und Datenbasis	45
3.1.1	Auswahl der Methode der Datenerhebung	45
3.1.2	Auswahl der Zielpopulation und der Stichprobe	48
3.2	Entwicklung der Messmodelle	51
3.2.1	Recherche von existierenden Items und Konstrukten	51
3.2.2	Entwicklung neuer Skalen gemäß C-OAR-SE	52

3.2.2.1	Definition und Erläuterung der Konstrukte	53
3.2.2.2	Objekt-Klassifizierung	55
3.2.2.3	Attribut-Klassifizierung	56
3.2.2.4	Rater-Klassifizierung	58
3.2.2.5	Skalenbildung	59
3.2.2.6	Enumeration	61
3.3	Kontrolle potentieller Messfehler	62
3.3.1	Nonresponse-Bias	62
3.3.2	Hindsight-Bias	63
3.3.3	Informant-Bias	65
3.3.4	Common-Method-Biases	66
3.3.4.1	Soziale Erwünschtheit und Impression-Management	68
3.3.4.2	Leniency-Bias	72
3.3.4.3	Reihenfolge-Bias	73
3.3.4.4	Acquiescence-Bias	74
3.4	Durchführung der empirischen Erhebungen	76
3.4.1	Vorstudien und Pretests	76
3.4.1.1	Vorstudie zur Entwicklung eines Impression-Management- Index	76
3.4.1.2	Pretest 1: Zuordnung der reflektiven Items	79
3.4.1.3	Pretest 2: Bewertung der reflektiven Items	84
3.4.1.4	Pretest 3: Bewertung durch Wissenschaftler und Praktiker mit Survey-Erfahrung	88
3.4.1.5	Pretest 4: Zeitdauer	89
3.4.1.6	Pretest 5: Kognitive Interviews	89
3.4.1.7	Pretest 6: Funktionalität des Web-Interfaces	90
3.4.1.8	Pretest 7: Großzahlige Erhebung zur Prüfung der Uni- dimensionalität	91
3.4.2	Durchführung der Hauptbefragung	92

4	Statistische Analysen und Diskussion der Ergebnisse	95
4.1	Datenaufbereitung und vorbereitende Analysen	95
4.1.1	Bestimmung und Diskussion der Rücklaufquote	95
4.1.1.1	Umgang mit fehlenden Daten	96
4.1.1.2	Ausschluss von Random-Respondern	100
4.1.1.3	Berechnung der Rücklaufquote	101
4.1.1.4	Beurteilung der Stichprobengröße	101
4.1.1.5	Suche nach fehlerhaften Werten	102
4.1.2	Prüfung auf Normalverteilung	103
4.1.3	Charakterisierung der Respondenten, Unternehmen und Projekte	107
4.2	Methodische Hintergründe zur kovarianzbasierten Analyse von Strukturgleichungsmodellen	110
4.2.1	Auswahl des statistischen Analyseverfahrens	110
4.2.2	Kriterien für die Gütebeurteilung der Messmodelle	112
4.2.2.1	Kriterien zur Prüfung der Reliabilität	112
4.2.2.2	Kriterien zur Prüfung der Unidimensionalität	114
4.2.2.3	Kriterien zur Prüfung der Validität	115
4.2.3	Kriterien für die Gütebeurteilung der Mess- und Strukturmodelle	117
4.3	Entwicklung und Beurteilung der Messmodelle	121
4.3.1	Operationalisierung der Projekteskalation	121
4.3.1.1	Verspätung des Projektabbruchs	124
4.3.1.2	Beurteilung des Ressourcen-Einsatzes	125
4.3.1.3	Investierter Anteil der geplanten Ressourcen	126
4.3.1.4	Diskussion der Projekteskalation als reflektives Konstrukt zweiter Ordnung	127
4.3.2	Operationalisierung der kognitiven Faktoren	130
4.3.2.1	Wahrnehmungsschwelle	130
4.3.2.2	Selektive Wahrnehmung	132
4.3.2.3	Self-Justification-Druck	133
4.3.2.4	Sunk-Cost-Effekt	134

4.3.2.5	Optimismus	134
4.3.2.6	Übermäßiger Optimismus	135
4.3.3	Operationalisierung der moderierenden Faktoren	137
4.3.3.1	Zeitpunkt des ersten Problemauftretens	137
4.3.3.2	Nicht hinterfragter Freiraum	138
4.3.3.3	Overconfidence	138
4.3.4	Operationalisierung der Variablen zur Prüfung der Common-Method-Biases	139
4.3.4.1	Method-Variance-Marker	139
4.3.4.2	Impression-Management-Index	140
4.3.4.3	Leniency-Index	141
4.3.4.4	Selektive Wahrnehmung anderer Manager	142
4.3.4.5	Freundschaft	143
4.3.5	Diskriminanzvalidität der Konstrukte	143
4.3.5.1	Konfirmatorische Faktorenanalyse	144
4.3.5.2	Fornell/Larcker-Kriterium	146
4.4	Analyse von Strukturmodellen der Projekteskalation	148
4.4.1	Analyse der kognitiven Faktoren	148
4.4.1.1	Eskalationsmodell ohne Kovarianzen zwischen den Faktoren (Strukturmodell 1)	148
4.4.1.2	Eskalationsmodell mit Kovarianzen zwischen den Faktoren (Strukturmodell 2)	150
4.4.1.3	Eskalationsmodell unter Berücksichtigung der kognitiven Dissonanz (Strukturmodell 3)	152
4.4.1.4	Exploratorische Modellselektion (Strukturmodell 4)	155
4.4.1.5	Zusammenführung der Ergebnisse der vier betrachteten Modelle	160
4.4.2	Analyse der moderierenden Effekte	162
4.4.2.1	Zeitpunkt des ersten Problemauftretens	163
4.4.2.2	Nicht hinterfragter Freiraum	164
4.4.2.3	Overconfidence	166

4.4.3	Beurteilung des Akteursmodells	166
4.5	Analyse der Maßnahmen des verhaltensorientierten Controllings	169
4.5.1	Wirkung der Maßnahmen auf die kognitiven Faktoren	169
4.5.1.1	Maßnahmen zur Reduzierung der Wahrnehmungsschwelle	170
4.5.1.2	Maßnahmen zur Reduzierung der selektiven Wahrnehmung	171
4.5.1.3	Maßnahmen zur Reduzierung des Self-Justification-Drucks	172
4.5.1.4	Maßnahmen zur Reduzierung des Sunk-Cost-Effekts	173
4.5.1.5	Maßnahmen zur Reduzierung des übermäßigen Optimismus	174
4.5.1.6	Maßnahmen zur Reduzierung des nicht hinterfragten Frei- raums	174
4.5.1.7	Zusammenfassende Diskussion der Einzelwirkungen der Maßnahmen	175
4.5.2	Wirkung aller Maßnahmen auf die Projekteskalation	176
4.5.3	Nutzung und Nutzen der Maßnahmen in der Unternehmenspraxis	178
4.6	Prüfung potentieller Fehlerquellen	182
4.6.1	Prüfung des Informant-Bias	182
4.6.2	Prüfung der Common-Method-Biases	185
4.6.2.1	Prüfung der Methodenvarianz mit der Marker-Variablen	185
4.6.2.2	Prüfung der sozialen Erwünschtheit und des Impression- Managements	187
4.6.2.3	Prüfung des Leniency-Bias	190
4.6.2.4	Prüfung des Reihenfolge-Bias	197
4.6.3	Prüfung weiterer Biases	198
4.6.4	Prüfung der Modellstabilität	200
5	Zusammenfassung und Fazit	201
5.1	Zentrale Ergebnisse	201
5.2	Implikationen für die Praxis	203
5.3	Implikationen für die Wissenschaft und weiterer Forschungsbedarf	205
5.4	Limitationen der Untersuchung	209

Anhang	213
A.1 Vorstudie zum Impression-Management: Einleitungstext	213
A.2 Sorting-Pretest: Begrüßungstext und Konstruktdefinitionen	216
A.3 Rating-Pretest: Begrüßungstext	218
A.4 Anschreiben	220
A.5 Fragebogen	222
A.6 Erstes Erinnerungsschreiben	232
A.7 Abschließendes Erinnerungsschreiben	234
A.8 Mittelwerte und Standardabweichungen	236
 Literaturverzeichnis	 241

Tabellenverzeichnis

2.1	Erweiterungen des ökonomischen Modells; in Anlehnung an Zayer (2007), S. 93	17
2.2	Großzahlige Untersuchungen in verwandten Themenfeldern	20
3.1	OAR-Definitionen von Projekteskalation und kognitiven Faktoren	54
3.2	Checkliste zur Formulierung von Items	57
3.2	Checkliste zur Formulierung von Items (Fortsetzung von Seite 57)	58
3.3	Elemente des Anschreibens	63
3.4	Maßnahmen zur Verringerung des Impression-Managements	71
3.5	Ergebnisse der Vorstudie zum Impression-Management	78
3.6	Ergebnisse des Sortierungspretests	81
3.6	Ergebnisse des Sortierungspretests (Fortsetzung von Seite 81)	82
3.6	Ergebnisse des Sortierungspretests (Fortsetzung von Seite 82)	83
3.7	Ergebnisse des Bewertungspretests	85
3.7	Ergebnisse des Bewertungspretests (Fortsetzung von Seite 85)	86
3.7	Ergebnisse des Bewertungspretests (Fortsetzung von Seite 86)	87
3.8	Ergebnisse des großzahligen Pretests	91
4.1	Standardfehler der Pfadkoeffizienten beim Bootstrap-Verfahren	106
4.2	Konfidenzintervalle der Pfadkoeffizienten beim Bootstrap-Verfahren	106

4.3	Gütebeurteilung von Messmodellen	117
4.4	Gütebeurteilung von Strukturmodellen	120
4.5	Gütekriterien des Messmodells „verspäteter Projektabbruch“	124
4.6	Gütekriterien des Messmodells „Beurteilung des Ressourcen-Einsatzes“	125
4.7	Gütekriterien des Messmodells „investierter Anteil der geplanten Ressourcen“	126
4.8	Gütekriterien des Messmodells zweiter Ordnung „Projekteskalation“	130
4.9	Gütekriterien des Messmodells „Wahrnehmungsschwelle“	131
4.10	Gütekriterien des Messmodells „selektive Wahrnehmung“	132
4.11	Gütekriterien des Messmodells „Self-Justification-Druck“	133
4.12	Gütekriterien des Messmodells „Sunk-Cost-Effekt“	134
4.13	Gütekriterien des Messmodells „Optimismus“	135
4.14	Variable „übermäßiger Optimismus“	136
4.15	Gütekriterien des Messmodells „Zeitpunkt des ersten Problemauftretens“	137
4.16	Gütekriterien des Messmodells „nicht hinterfragter Freiraum“	138
4.17	Gütekriterien des Messmodells „Overconfidence“	139
4.18	Gütekriterien des Messmodells „Kapitalmarktorientierung“	140
4.19	Mittelwerte und Standardabweichungen des Index „Impression-Management“	140
4.20	Mittelwerte und Standardabweichungen des Index „Leniency“	141
4.21	Gütekriterien des Messmodells „selektive Wahrnehmung anderer Manager“	142
4.22	Gütekriterien des Messmodells „Freundschaft“	143
4.23	Unterscheidung zwischen signifikanten und nicht signifikanten Korrelationen zwischen den Konstrukten	145
4.24	Diskriminanzvalidität anhand des Fornell/Larcker-Kriteriums	147
4.25	Empirische Beurteilung der Erweiterungen des ökonomischen Modells	168
4.26	Maßnahmen zur Überwindung der Wahrnehmungsschwelle	171

4.27 Maßnahmen zur Überwindung der selektiven Wahrnehmung	172
4.28 Maßnahmen zur Reduzierung des Self-Justification-Drucks	172
4.29 Maßnahmen zur Reduzierung des Sunk-Cost-Effekts	173
4.30 Maßnahmen zur Reduzierung des übermäßigen Optimismus	174
4.31 Maßnahmen zur Reduzierung des nicht hinterfragten Freiraums	175
A.1 Impression-Management-Subscale	214
A.2 Latente Konstrukte	236
A.2 Latente Konstrukte (Fortsetzung von Seite 236)	237
A.2 Latente Konstrukte (Fortsetzung von Seite 237)	238
A.3 Maßnahmen	239
A.3 Maßnahmen (Fortsetzung von Seite 239)	240

Abbildungsverzeichnis

1.1	Aufbau der Arbeit	8
2.1	Ausgewählte Publikationen zur Eskalation des Commitments	13
2.2	Gesamtmodell der Eskalationsfaktoren, Quelle: Zayer 2007, S. 225	22
2.3	Eskalationsmodell für die empirische Untersuchung; in Anlehnung an Zayer 2007, S. 213 (vereinfacht und leicht modifiziert)	28
3.1	C-OAR-SE-Prozedur; in Anlehnung an Rossiter 2002, S. 306	53
4.1	Charakteristika der Projekte	108
4.2	Charakteristika der Umfrageteilnehmer	109
4.3	Projekteskalation als reflektives Konstrukt zweiter Ordnung	128
4.4	Modell ohne Kovarianzen zwischen den kognitiven Faktoren (Modell 1)	149
4.5	Modell mit Kovarianzen zwischen den kognitiven Faktoren (Modell 2)	150
4.6	Modell mit kognitiver Dissonanz (Modell 3)	155
4.7	Specification Search der Wirkungen der kognitiven Faktoren	157
4.8	Eines von sieben exploratorisch ermittelten Modellen (Modell 4)	159
4.9	Zeitpunkt des ersten Problemauftretens als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)	164
4.10	Nicht hinterfragter Freiraum als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)	165

4.11	Overconfidence als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)	167
4.12	Wirkung aller Maßnahmen	177
4.13	Vergleich der Häufigkeit der Maßnahmenverwendung	180
4.14	Investierte Ressourcen und Fertigstellungsgrad	181
4.15	Gruppenvergleich Entscheidungsverantwortliche vs. Nicht-Entscheidungsverantwortliche für Modell 2	183
4.16	Eskalation (Modell 2) ohne Impression-Management-Index	188
4.17	Eskalation (Modell 2) mit Impression-Management-Index	189
4.18	Eskalation (Modell 2), beurteilt durch Nicht-Entscheidungsverantwortliche	190
4.19	Eskalation (Modell 2) mit Leniency-Index	191
4.20	Leniency als moderierender Faktor für Modell 2	192
4.21	Eskalation (Modell 2) mit Freundschafts-Konstrukt	194
4.22	Freundschaft als moderierender Faktor für Modell 2	196

Abkürzungsverzeichnis

AOS	Accounting, Organizations and Society (Zeitschrift)
BCC	Browne-Cudeck-Criterion
BRIA	Behavioral Research in Accounting (Zeitschrift)
CFI	Comparative-Fit-Index
DEV	Durchschnittlich erfasste Varianz
df	Degrees of Freedom
EM	Expectation Maximization
F&E	Forschung und Entwicklung
GLS	Generalized Least Squares
H	Hypothese
HaPI	Health and Psychosocial Instruments (Datenbank psychologischer Tests)
JAR	Journal of Accounting Research (Zeitschrift)
MANOVA	Multivariate Analysis of Variance
MAR	Missing at Random
MCAR	Missing Completely at Random
ML	Maximum Likelihood
MTMM	Multitrait-Multimethod
MV	Method Variance
NMAR	Not Missing at Random
NNFI	Nonnormed-Fit-Index
OAR	Object, Attribute, Rater
PLS	Partial Least Square
RMSEA	Root Mean Square Error of Approximation
SRMR	Standardized Root Mean Square Residual
TAR	The Accounting Review (Zeitschrift)
TLI	Tucker-Lewis-Index

1 Einleitung

1.1 Das Phänomen der Eskalation des Commitments

Im Mittelpunkt der vorliegenden Arbeit steht die Eskalation von Projekten.¹ Der Begriff Eskalation bezieht sich dabei auf das Phänomen der *Eskalation des Commitments*, worunter die menschliche Eigenart verstanden wird, irrational lange an scheiternden Verhaltensweisen festzuhalten.² So führen Menschen Projekte häufig auch dann weiter, wenn die verfügbaren Informationen für einen Abbruch sprechen.³ Scheiternde Projekte werden dadurch nicht zum ökonomisch rationalen Zeitpunkt, sondern systematisch zu spät beendet.⁴ Im betriebswirtschaftlichen Kontext führt die Eskalation des Commitments zu beträchtlichen Fehlallokationen von Ressourcen⁵ und zu organisationalen Störungen.⁶ Meredith stellt deshalb fest: “Yet, the importance of quickly terminating a failing project can hardly be overemphasized.”⁷

¹ Für die Abgrenzung des Geltungsbereichs vgl. Abschnitt 2.2.2.1 auf Seite 23.

² Für eine detailliertere Betrachtung der Begriffe Eskalation des Commitments (Escalation of Commitment) und Rationalität siehe Abschnitt 2.1.2 auf Seite 12.

³ Vgl. exemplarisch Barton/Duchon/Dunegan 1989, Schmidt/Calantone 2002, S. 104, Staw 1976, Staw 1997 und Staw/Ross 1987a.

⁴ Der Begriff systematisch bezieht sich dabei auf den Umstand, dass Projekte nicht normalverteilt zu früh oder zu spät abgebrochen werden, sondern mehr Projekte zu spät als zu früh abgebrochen werden; vgl. Cheng et alii 2003, S. 64, Lange 1993, S. 147 ff., und Meredith 1988, S. 31.

⁵ Vgl. Lange 1993, S. 79 ff.

⁶ Vgl. Meredith 1988, S. 31.

⁷ Meredith 1988, S. 31.

Die Eskalation von Projekten wird in einer Vielzahl von Branchen und für unterschiedliche Situationen berichtet; so unter anderem in der Forschung und Entwicklung (F&E),⁸ bei der Entwicklung neuer Produkte,⁹ bei Softwareprojekten,¹⁰ bei Personalentscheidungen,¹¹ bei großen Bauprojekten,¹² bei der Vergabe von Risikokapital¹³ sowie bei strategischen Entscheidungen¹⁴ und bei Markteintrittsprojekten.¹⁵ Die aufgeführten Studien zeigen, dass die Eskalation des Commitments (nach wie vor) ein Problem von hoher Relevanz darstellt. Obwohl sich die Forschung schon seit über 30 Jahren mit übermäßigem Commitment befasst,¹⁶ wurde das Problem weder aus praktischer Sicht zufriedenstellend gelöst noch aus wissenschaftlicher Perspektive vollständig erklärt. Die vorliegende Arbeit verfolgt deshalb das Ziel, zu einem besseren Verständnis des Phänomens der Eskalation des Commitments beizutragen.

1.2 Relevanz der Arbeit

1.2.1 Relevanz für die Wissenschaft

Verschiedene Kriterien können herangezogen werden, um die wissenschaftliche Relevanz zu prüfen. Maines/Salamon/Sprinkle heben beispielsweise zwei Aspekte hervor, die einen begründeten Ansatz für wissenschaftliche Forschung bilden. Zum einen kann der Forscher eine Theorie heranziehen oder entwickeln, die relevant für das entsprechende Entscheidungsproblem ist, bisher jedoch noch nicht berücksichtigt wurde. Zum anderen können die Methoden hinterfragt werden, mit denen die Erkenntnisse generiert wurden, auf denen die gegenwärtige Ansicht

⁸ Vgl. Lange 1993 und Zayer 2007.

⁹ Vgl. Schmidt/Calantone 2002.

¹⁰ Vgl. Heng/Tan/Wei 2003, Keil/Rai/Mann 2000, Keil/Rai/Mann/Zhang 2003, Keil/Robey 1999, Keil/Robey 2001, Keil et alii 1995, Keil et alii 2000, Montealegre/Keil 2000, Smith/Keil 2003, Smith/Keil/Depledge 2001, Tan et alii 2003, Tiwana/Keil/Fichman 2006 und Zhang et alii 2003.

¹¹ Vgl. Bazerman/Beekun/Schoorman 1982 und Schoorman 1988.

¹² Vgl. Ross/Staw 1993.

¹³ Vgl. Guler 2007.

¹⁴ Vgl. Kisfalvi 2000 und Lant/Hurley 1999.

¹⁵ Vgl. Camerer 1999, Rapoport 1995 und Rapoport et alii 1998.

¹⁶ Vgl. Abschnitt 2.1.2 auf Seite 12.

basiert.¹⁷ Die wissenschaftliche Motivation dieser Arbeit betrifft beide Argumentationspunkte, wie im Folgenden näher ausgeführt wird.

Die Wissenschaft beschäftigt sich mit dem Phänomen der Eskalation des Commitments seit über 30 Jahren.¹⁸ Dabei leisten insbesondere verhaltenswissenschaftliche Erkenntnisse einen wichtigen Beitrag zum Verständnis der Eskalation. Im Verlauf der Jahrzehnte wurden mehr und mehr kognitive Faktoren identifiziert, die sich auf die Eskalation auswirken.¹⁹ Die Bezeichnung „kognitiv“ umfasst Prozesse der Wahrnehmung, Aufmerksamkeit und Erinnerung²⁰ sowie das Wissen, die Meinung oder die Auffassung über die Umgebung, über sich selbst oder über das eigene Verhalten.²¹

Die Forschung fand dabei in der Regel in Form von Laborexperimenten statt.²² Mit dieser Forschungsmethode sind zwei Restriktionen verbunden. Zum einen kann mit Experimenten nur eine stark begrenzte Anzahl von Variablen simultan untersucht werden, da andernfalls die Anzahl der notwendigen Gruppen zu stark ansteigt. Zum anderen stellt sich bei Experimenten die Frage nach der externen Validität, d. h., der Übertragbarkeit auf Situationen in der realen Unternehmenswelt.

Korrespondierend mit diesen beiden Restriktionen finden sich zwei Forderungen in der Eskalationsliteratur. Einerseits wird ein starker Bedarf an Studien konstatiert, die mehrere kognitive Faktoren simultan untersuchen.²³ Zum anderen wird hervorgehoben, dass einer Prüfung der bisherigen Forschungsergebnisse im realen Unternehmenskontext eine hohe Bedeutung zukommt. Ein großer Teil der Eskalationsstudien verwendet Laborexperimente mit *Studenten* der Wirtschaftswissenschaften als Subjekte.²⁴ Da Erfahrung einen signifikanten Einfluss auf die Eskalationstendenz eines Individuums hat,²⁵ warnen beispielsweise Salter/Sharp davor, die Erkennt-

¹⁷ Vgl. Maines/Salamon/Sprinkle 2006, S. 91.

¹⁸ Für eine frühe Arbeit zum Commitment vgl. Kiesler/Sakumura 1966, für die Thematik verspäteter Projektabbrüche vgl. Jones/McLean 1970, S. 8.

¹⁹ Vgl. Abbildung 2.1 auf Seite 13.

²⁰ Vgl. Kellogg 2002, S. 55.

²¹ Vgl. Festinger 1957, S. 3.

²² Vgl. Simonson/Staw 1992, S. 421.

²³ Vgl. Camerer 1999, S. 81, Keil et alii 2003, S. 2 und 260. Sehr explizit beispielsweise das folgende Zitat von Keil/Rai/Mann: "Our findings suggest that escalation is a complex phenomenon and that more sophisticated models are needed to explain behavior."; Keil/Rai/Mann 2000, S. 656.

²⁴ Vgl. Simonson/Staw 1992, S. 421.

²⁵ Die Ausbildung und Erfahrung kann auf verschiedene Weisen das Eskalationsverhalten beeinflussen, beispielsweise durch eine Senkung des Confirmation Bias (vgl. Cloyd/Spilker 2000) und der Overconfidence

nisse aus Laborexperimenten mit unerfahrenen Studenten auf Manager und ihr reales Handeln zu übertragen.²⁶

Neben der Frage nach der Übertragbarkeit der Ergebnisse im Hinblick auf die spezifischen Eigenschaften der untersuchten *Personen* ist bei Experimenten die externe Validität auch durch die Art der untersuchten *Problemsituation* gefährdet. Es besteht das Risiko, dass sich die Personen in komplexen Problemsituationen der realen Welt anders verhalten als bei kurzen und verhältnismäßig einfachen Experimenten.²⁷ Exemplarisch für die Problematik der externen Validität in der Eskalationsforschung sei folgendes Zitat angeführt, das in den abschließenden Bemerkungen einer experimentellen Studie zur Eskalation des Commitments zu finden ist: “Because our primary aim was to disentangle previously confounded variables, we sought to maximize internal validity [. . .] Nevertheless, the precision gained in the present paradigm comes at the expense of external validity. It is thus imperative that field research be conducted to assess the generalizability of our findings.”²⁸

Von hoher wissenschaftlicher Relevanz sind demzufolge zum gegenwärtigen Stand der Forschung²⁹ insbesondere großzahlige Befragungen, die es ermöglichen, eine verhältnismäßig *große Anzahl* von Variablen im *realen Unternehmenskontext* zu prüfen.³⁰ Damit kann den bisher überwiegend verwendeten Experimenten eine andere Forschungsmethode zur Seite gestellt werden. Darüber hinaus bietet eine solche Untersuchung die Chance, die bestehenden Theorien zu verfeinern und beispielsweise ihre Kontextabhängigkeit zu untersuchen. Insgesamt ist als Zwischenergebnis festzuhalten, dass aus wissenschaftlicher Sicht begründeter Bedarf für eine weitere Beschäftigung mit der Thematik der Eskalation des Commitments besteht.

(vgl. Russo/Schoemaker 1992, S. 10). Eine experimentelle Untersuchung von Salter/Sharp kommt zu dem Ergebnis, dass die Eskalationstendenz mit zunehmender Erfahrung des Managers geringer wird (vgl. Salter/Sharp 2001, S. 40). Gleichzeitig gibt es jedoch auch empirische Befunde, dass die Eskalationstendenz stärker ist, wenn das Individuum in einem Gebiet entscheidet, das zu seinem Fachgebiet gehört (vgl. Fox/Schmid/Yinon 1995, S. 255).

²⁶ “Our confirmation of the effect of experience suggests that the use of inexperienced student subjects in escalation of commitment research as proxies for managers may overestimate the willingness of managers to escalate. Therefore, the findings of studies in the psychology literature [. . .] using undergraduate student subjects may not be generalizable to manager populations.”; Salter/Sharp 2001, S. 41.

²⁷ Vgl. Salter/Sharp 2001, S. 41.

²⁸ Bobocel/Meyer 1994, S. 363.

²⁹ Für eine Erläuterung des Standes der Forschung vgl. Abschnitt 2.1.2 auf Seite 12.

³⁰ Die Diskussion zur Auswahl der Methode wird in Abschnitt 3.1.1 auf Seite 45 wieder aufgegriffen und fortgeführt.

1.2.2 Relevanz für die Praxis

Scheiternde Projekte, die einen Abbruch oder zumindest einen Neustart erfordern, sind in vielen Branchen an der Tagesordnung. Um ihre Größenordnung abschätzen zu können, seien hier exemplarisch einige Zahlen aus den zwei Bereichen (1) Informationssysteme und (2) F&E genannt.

(1) Etwa 50 Prozent aller Data-Warehousing-Projekte scheitern.³¹ 80 Prozent aller Knowledge-Management-Systeme liefern keine Resultate;³² die dabei verursachten Kosten belaufen sich auf etliche Milliarden Dollar. Im Jahr 1995 wurden beispielsweise allein im Zusammenhang mit Informationssystemen 81 Mrd. US-Dollar für abgebrochene Softwareprojekte ausgegeben.³³ Für das Jahr 1998 werden die Kosten für gescheiterte IT-Projekte in den USA auf 75 Mrd. US-Dollar geschätzt.³⁴

(2) Je nach Branche betragen die Aufwendungen für F&E bis zu zehn Prozent vom Umsatz, wie beispielsweise in der Elektroindustrie.³⁵ Allen/Norris ermitteln für *abgebrochene* F&E-Projekte eine durchschnittliche Überschreitung der geplanten Projektdauer um 76 Prozent,³⁶ was die Schlussfolgerung nahe legt, dass Abbruchentscheidungen häufig erst dann als Handlungsalternative erwogen werden, wenn die ursprünglich geplante Gesamtdauer bereits überschritten ist.³⁷ Dies ist umso wichtiger, als bei risikoreichen Investitionsprojekten (wie beispielsweise in der F&E) *erfolgreiche* Projekte eher die Ausnahme als die Regel sind.³⁸ Nach Abschätzungen von Commes/Lienert erreichen nur 31 von 100 F&E-Projekten die Phase der Markteinführung und nur 12 sind wirtschaftlich erfolgreich.³⁹ Eine Betrachtung der damit verbundenen Kosten zeigt, dass mehr Entwicklungskosten in abgebrochene oder erfolglose Projekte fließen als in erfolgreiche.⁴⁰ Dies allein sagt noch nichts über rationale oder irrationale Entscheidungen aus,

³¹ Vgl. Watson/Haley 1997.

³² Vgl. Tiwana 2000.

³³ Vgl. Johnson 1995.

³⁴ Vgl. Johnson 2000.

³⁵ Vgl. Schmeisser et alii 2006, S. 60.

³⁶ Vgl. Allen/Norris 1970, S. 278. Für die Projekte in der vorliegenden Studie beträgt der durchschnittlich investierte Anteil der Mannmonate 116 Prozent des ursprünglich geplanten Wertes; vgl. Abbildung 4.14 auf Seite 181.

³⁷ Vgl. Lange 1993, S. 76.

³⁸ Vgl. beispielsweise Booz-Allen-Hamilton-Inc. (Hrsg.) 1968, S. 11, Lange 1993, S. 78, und Mansfield et alii 1971, S. 52.

³⁹ Vgl. Commes/Lienert 1983, S. 349.

⁴⁰ Vgl. beispielsweise Lange 1993, S. 79 ff.

sondern spiegelt zunächst nur das unternehmerische Risiko wider. In Verbindung mit der Erkenntnis, dass Projekte systematisch zu spät abgebrochen werden, geben diese Daten jedoch einen ungefähren Anhaltspunkt für das Ausmaß des wirtschaftlichen Schadens, der durch die Eskalation des Commitments droht.

Frühzeitige Abbrüche erfolgloser Projekte sparen Unternehmen Zeit und Kosten und reduzieren organisationale Störungen, die durch scheiternde Projekte verursacht werden.⁴¹ Empirische Studien zeigen jedoch, dass Projekte mit geringen Erfolgsaussichten in der Praxis häufig zu spät abgebrochen werden.⁴²

Während sich die Forschung seit langem mit den Faktoren beschäftigt, die zur Eskalation eines Projekts führen, findet eine konsequente Untersuchung deeskalierender Maßnahmen erst seit wenigen Jahren statt.⁴³ In der Literatur wird gefordert, die Deeskalationsforschung auszuweiten, da sie das Potential bietet, Managern bei der Überwindung der Eskalation des Commitments zu helfen.⁴⁴ Die vorliegende Studie greift diese Forderung auf, indem sie nicht nur die kognitiven Faktoren untersucht,⁴⁵ die zur Eskalation von Projekten beitragen, sondern darüber hinaus auch die Wirksamkeit verschiedener Maßnahmen des verhaltensorientierten Controllings zur Reduzierung der Eskalation prüft.⁴⁶

Die Studie weist somit in zweierlei Hinsicht eine hohe Relevanz für die Praxis auf. Zum einen kann ein besseres Verständnis der relativen Einflussstärke der *kognitiven Faktoren* helfen, Prioritäten bei der Bekämpfung der Eskalation zu setzen. Zum anderen können die Analysen aufzeigen, welche *Maßnahmen* sich im realen Unternehmenskontext als besonders wirksam erweisen.

⁴¹ Vgl. Meredith 1988, S. 31.

⁴² Vgl. beispielsweise Lange 1993, S. 147 ff. Meredith stellt explizit fest, "it has too frequently been the case that projects were allowed to continue well past the time when it was obvious they were headed for disaster."; Meredith 1988, S. 31.

⁴³ Keil/Robey 1999, S. 63, und Montealegre/Keil 2000, S. 419.

⁴⁴ Vgl. Simonson/Staw 1992, S. 419.

⁴⁵ Für eine Definition des Begriffs *Kognition* siehe Abschnitt 2.1.2 auf Seite 12.

⁴⁶ Vgl. Abschnitt 4.5 auf Seite 169.

1.3 Zielsetzung der Arbeit

Aus den geschilderten, für Wissenschaft und Praxis relevanten Forschungslücken lassen sich zwei Forschungsfragen ableiten. Die erste Forschungsfrage bezieht sich auf das bessere Verständnis der kognitiven Eskalationsfaktoren in der Unternehmenspraxis.

Forschungsfrage 1: Lässt sich die Wirkung der theoretisch und experimentell ermittelten kognitiven Faktoren auf die Projekteskalation empirisch im realen Unternehmenskontext bestätigen? Wenn ja, welche relative Einflussstärke weisen die verschiedenen kognitiven Faktoren auf die Projekteskalation auf?

Die zweite Forschungsfrage zielt auf ein besseres Verständnis von Prozeduren zur Deeskalation ab. Dazu sollen verschiedene Maßnahmen des verhaltensorientierten Controllings im Hinblick auf ihre Wirksamkeit untersucht werden.

Forschungsfrage 2: Lässt sich die Wirkung der verhaltenswissenschaftlich basierten Maßnahmen auf die kognitiven Faktoren empirisch im Unternehmenskontext bestätigen? Wenn ja, welche relative Einflussstärke weisen die verschiedenen Maßnahmen auf die kognitiven Faktoren auf?

1.4 Aufbau der Arbeit

Die vorliegende Arbeit ist in sechs Kapitel untergliedert. Abbildung 1.1 auf der nächsten Seite gibt einen Überblick über den Aufbau der Arbeit.

Kapitel 1 stellt das Konzept der Eskalation des Commitments vor und erläutert Relevanz, Zielsetzung und Aufbau der Arbeit.

In Kapitel 2 wird die theoretische Basis für die Untersuchung entwickelt. Dazu werden konzeptionelle Grundlagen zur Eskalation von Projekten darstellt. Im Anschluss daran wird das Modell des Zusammenwirkens der kognitiven Faktoren und der Maßnahmen des verhaltensorientierten Controllings hergeleitet.

Kapitel 3 stellt die Methodik und die Durchführung der empirischen Erhebung vor. Zu diesem Zweck werden zunächst die Datenerhebungsmethode und die Datenbasis diskutiert. Im

1 Einleitung		
1.1 Das Phänomen der Eskalation des Commitments		
1.2 Relevanz der Arbeit	1.3 Zielsetzung der Arbeit	1.4 Aufbau der Arbeit
2 Theoretische Basis der Untersuchung		
2.1 Konzeptionelle Grundlagen		
2.2 Herleitung des Modells	2.3 Erläuterung der Variablen und Herleitung der Hypothesen	
3 Methodik und Durchführung der empirischen Studie		
3.1 Diskussion der Datenerhebung und Datenbasis	3.2 Entwicklung der Messmodelle	
3.3 Kontrolle potentieller Messfehler	3.4 Durchführung der empirischen Erhebungen	
4 Statistische Analysen und Diskussion der Ergebnisse		
4.1 Datenaufbereitung und vorbereitende Analysen		
4.2 Methodische Hintergründe zur kovarianzbasierten Analyse von Strukturgleichungsmodellen		
4.3 Entwicklung und Beurteilung der Messmodelle		
4.4 Analyse von Strukturmodellen der Projekteskalation		
4.5 Analyse der Maßnahmen des verhaltensorientierten Controllings		
4.6 Prüfung potentieller Fehlerquellen		
5 Zusammenfassung und Fazit		
5.1 Zentrale Ergebnisse	5.2 Implikationen für die Praxis	
5.3 Implikationen für die Wissenschaft und weiterer Forschungsbedarf		5.4 Limitationen der Untersuchung

Abbildung 1.1: Aufbau der Arbeit

Anschluss daran wird die Prozedur zur Entwicklung neuer Messmodelle skizziert. Darüber hinaus wird erläutert, wie die Kontrolle potentieller Messfehler beim Design der Befragung berücksichtigt wurde. Abschließend werden die verschiedenen Pretest-Stufen und die Durchführung der Hauptbefragung vorgestellt.

In Kapitel 4 werden die statistischen Analysen präsentiert und die Ergebnisse diskutiert. Dabei wird zuerst die Datenaufbereitung geschildert. Danach werden die methodischen Hintergründe zur kovarianzbasierten Analyse von Strukturgleichungsmodellen beschrieben. Als Kern der Arbeit werden die Messmodelle beurteilt und die Hypothesen in Bezug auf die kognitiven Faktoren, die moderierenden Effekte und die Maßnahmen des verhaltensorientierten Controllings geprüft. Zusätzlich werden die Ergebnisse kritisch hinterfragt und explizit auf potentielle Messfehler und systematische Verzerrungen hin untersucht.

Kapitel 5 fasst die Ergebnisse zusammen und diskutiert Limitationen der Untersuchung sowie die Implikationen für Praxis und Wissenschaft.

2 Theoretische Basis der Untersuchung

2.1 Konzeptionelle Grundlagen

2.1.1 Entwicklung der Behavioral-Accounting-Forschung und Einordnung der Arbeit

Unter den verschiedenen wissenschaftlichen Forschungszweigen ist die vorliegende Arbeit den Sozialwissenschaften und unter diesen wiederum der Ökonomik zuzuordnen. Die wissenschaftliche Literatur bietet verschiedene Definitionen von Ökonomik.¹ Eine verbreitete Definition stammt von Robbins: „Economics is the science which studies human behavior as a relationship between ends and scarce means which have alternative uses.“²

Trotz verschiedener Definitionen herrscht weitgehend Einverständnis darüber, dass die Ökonomik eine *positive* Wissenschaft ist und als Erfahrungswissenschaft versucht, die Phänomene der sozialen Welt zu verstehen und nach Möglichkeit zu prognostizieren. Strittig hingegen ist die Frage, ob die Ökonomik Empfehlungen zur Gestaltung der sozialen Welt geben kann und sollte, also *normativ* vorgehen sollte.³

Die vorliegende Arbeit folgt dem Verständnis von Homann/Suchanek, die (ähnlich den Klassikern der ökonomischen Theorie von Adam Smith bis Oliver E. Williamson) die normative Seite

¹ Vgl. Homann/Suchanek 2000, S. 2.

² Robbins 1935, S. 15. Bei dieser Definition verweist Robbins auf Mengers „Grundsätze der Volkswirtschaftslehre“, Mises „Die Gemeinwirtschaft“, Fetters „Economic Principles“, Striegls „Die ökonomischen Kategorien und die Organisation der Wirtschaft“ und Mayers Aufsatz „Untersuchungen zu dem Grundgesetz der wirtschaftlichen Wertrechnung“ in der Zeitschrift für Volkswirtschaft und Sozialpolitik. Für eine Kritik an dieser Definition vgl. beispielsweise Homann/Suchanek 2000.

³ Vgl. Homann/Suchanek 2000, S. 27.

der Ökonomik nicht nur unterstützen, sondern als Ziel auch der positiven Analysen betrachten. Ökonomische Forschung und Theoriebildung folgen damit dem Grundsatz der „*Erklärung zwecks Gestaltung*“.⁴ Dieser Anspruch deckt sich mit dem spezifischen Untersuchungsgegenstand dieser Arbeit, dem verhaltensorientierten Controlling. Denn Controllingsysteme bilden die Realität nicht nur ab, sondern gestalten sie auch mit.⁵

Das in der deutschsprachigen Begriffswelt als verhaltensorientiertes Controlling bezeichnete Forschungsfeld⁶ lässt sich auf internationaler Ebene am ehesten dem Behavioral Accounting zuordnen. In der Literatur existiert allerdings kein eindeutiges Verständnis des Begriffs Behavioral Accounting.⁷ Eine einflussreiche Definition stammt von Hofstedt/Kinard: “Behavioral accounting research may be defined as the study of the behavior of accountants or the behavior of non-accountants as they are influenced by accounting functions and reports.”⁸ Diesem Verständnis folgt die vorliegende Arbeit.

Die Wurzeln des Behavioral Accounting liegen im Management Accounting.⁹ Dieses entwickelte sich in den frühen 1950er Jahren als Resultat der wissenschaftlichen Beschäftigung mit dem Thema Management und verfolgte das Ziel, den Profit der Unternehmen durch eine Steigerung von Effizienz und Effektivität weiter zu erhöhen.¹⁰ Um dazu beizutragen, stellen Simon et alii die Forderungen an das „Controller’s Department“, Informationen hoher Qualität zu minimalen Kosten bereitzustellen und die langfristige Entwicklung kompetenter Führungskräfte im Accounting und in der Geschäftsführung zu fördern.¹¹

Gegen Mitte und Ende der 1950er Jahre gerieten die traditionellen Annahmen der Ökonomik, wie beispielsweise der homo oeconomicus als vollständig rationaler Nutzenmaximierer, zunehmend in die Kritik.¹² Simon brachte in die wissenschaftliche Diskussion das Konzept der

⁴ Homann/Suchanek 2000, S. 28; im Original fett und kursiv hervorgehoben.

⁵ Vgl. für diese Aussage, bezogen auf *Accountingsysteme*, Morgan 1988, S. 482. Für praxisorientierte Gestaltungsempfehlungen in dieser Arbeit vgl. Abschnitt 4.5.3 auf Seite 178 und Abschnitt 5.2 auf Seite 203.

⁶ Vgl. beispielsweise Hirsch 2007, Meyer 2008 und Weber et alii 2003. Analog dazu steht das kognitionsorientierte Controlling dem Cognitive Accounting nahe, das als Teilgebiet des Behavioral Accounting aufgefasst werden kann; vgl. Gerling 2007, S. 114. Zum kognitionsorientierten Controlling vgl. Lingnau 2006.

⁷ Vgl. Lord 1989, S. 124.

⁸ Hofstedt/Kinard 1970, S. 43.

⁹ Vgl. beispielsweise Birnberg/Shields 1989, S. 27.

¹⁰ Vgl. beispielsweise Simon et alii 1954, S. 1.

¹¹ Vgl. Simon et alii 1954, S. 1.

¹² Vgl. beispielsweise Caplan 1966.

Bounded Rationality ein.¹³ In diesem Konzept wird die Annahme eines vollständig rationalen Individuums als unzutreffend verworfen und stattdessen begrenzte Rationalität angenommen. Auf Grund der limitierten intellektuellen Kapazitäten im Vergleich zum Komplexitätsgrad der Probleme, mit denen Individuen und Organisationen konfrontiert werden, erfordert rationales Verhalten vereinfachte Modelle, die sich auf einige wesentliche Merkmale des Problems beschränken.¹⁴ Im Zuge dieser Entwicklungen entstand ein neues Modell zur Erklärung von unternehmensbezogenen Fragestellungen, die „Behavioral Theory of the Firm“,¹⁵ in dem Manager nun als Akteure mit begrenzter Rationalität (*bounded rationality*) angesehen wurden.

Die geschilderte Entwicklung ging auch an der Accounting-Forschung nicht spurlos vorüber.¹⁶ Mit Beginn der 1960er Jahre wurden Vorwürfe laut, die Accounting-Forschung nehme die Erkenntnisse der Verhaltenswissenschaften nicht ausreichend zur Kenntnis.¹⁷ Zwar waren schon lange zuvor vereinzelt Gedanken zu verhaltensbezogenen Aspekten in der Accounting-Literatur erwähnt worden,¹⁸ doch erst ab 1960 erschienen regelmäßig Forschungsbeiträge zu diesem Thema.¹⁹ 1964 veröffentlichte das *Journal of Accounting Research* (JAR) als erste akademische Accounting-Zeitschrift einen *empirischen* Artikel zum Behavioral Accounting:²⁰ „The Effects of Alternative Accounting Techniques on Certain Management Decisions“ von Dyckman.²¹

Einen wichtigen Meilenstein in der Behavioral-Accounting-Forschung stellen die Arbeiten von Lindhe, Livingstone und Sorter et alii dar.²² War man zuvor davon ausgegangen, dass das Management rational entscheide, lieferten diese Untersuchungen zur Nutzung von Abschreibungsverfahren Evidenzen dafür, dass die Entscheidungen vielmehr von individuellen Charakteristiken geprägt sind. Livingstone und Sorter et alii führten dies auf psychologische Faktoren zurück.²³

¹³ Vgl. Simon 1957.

¹⁴ Vgl. March/Simon 1958, S. 169.

¹⁵ Vgl. Cyert/March 1963.

¹⁶ Für eine Analyse der frühen *empirischen* Behavioral-Accounting-Forschung vgl. Birnberg 1973, für eine umfassende Betrachtung der Behavioral-Accounting-Forschung von 1952 bis 1981 vgl. Lord 1989.

¹⁷ Vgl. beispielsweise Devine 1960. Ebenso Birnberg/Nath 1967, die betonten: „If accounting research is going to assist in the development of better managerial accounting techniques, effort must be made to utilize findings from the behavioral sciences.“; Birnberg/Nath 1967, S. 479.

¹⁸ Vgl. beispielsweise Dent 1931 und Ferguson 1920; beide Quellen zitiert nach Lord 1989, S. 126.

¹⁹ Vgl. Lord 1989, S. 127.

²⁰ Vgl. Lord 1989, S. 129.

²¹ Vgl. Dyckman 1964.

²² Vgl. Lindhe 1963, Livingstone 1967 und Sorter et alii 1966.

²³ Vgl. Livingstone 1967 und Sorter et alii 1966.

In den 1970er Jahren stieg das Interesse an den verhaltenswissenschaftlichen Aspekten im Accounting stark an. Swieringa/Weik listen über 100 Experimente auf, die zwischen 1970 und 1981 in den Zeitschriften Accounting, Organizations and Society (AOS), The Accounting Review (TAR) und Journal of Accounting Research (JAR) publiziert wurden.²⁴ Diese Entwicklung setzte sich in den 1980er Jahren fort.²⁵

Seit den 1980er Jahren kann die Behavioral-Accounting-Forschung als reifes und akzeptiertes Feld in der Accounting-Forschung betrachtet werden.²⁶ Heute gehört das *Behavioral Accounting* zu den Haupt-Forschungszweigen des Accounting.²⁷ Innerhalb dieses Feldes stellt das Phänomen der Eskalation des Commitments ein Thema dar, das bis heute Interesse in der Forschung findet.²⁸

2.1.2 Begriffsklärungen und Fokus der Arbeit

Der Begriff des Commitments stammt aus der Psychologie und wird dort als „binding of the individual to behavioral acts“²⁹ definiert. Unter einer *Eskalation* des Commitments ist demnach die *übermäßige* Bindung an ein Verhalten zu verstehen. Diese wird durch verschiedene kognitive Faktoren bedingt,³⁰ deren Einfluss in der vorliegenden Arbeit untersucht wird.

Im Laufe der letzten vier Jahrzehnte wurde eine Vielzahl von kognitiven Faktoren untersucht, die im Zusammenhang mit verspäteten Abbruchentscheidungen eine Rolle spielen können. Abbildung 2.1 auf der nächsten Seite gibt einen Überblick über die zeitliche Entwicklung der kognitiven Erklärungsansätze, die im weiteren Verlauf der Arbeit näher untersucht werden: Self-Justification, Wahrnehmungsschwelle, Sunk-Cost-Effekt, Optimismus, selektive Wahrnehmung und Overconfidence.³¹

²⁴ Vgl. Swieringa/Weik 1982, S. 128.

²⁵ Vgl. Bamber 1993.

²⁶ Vgl. Lord 1989, S. 139.

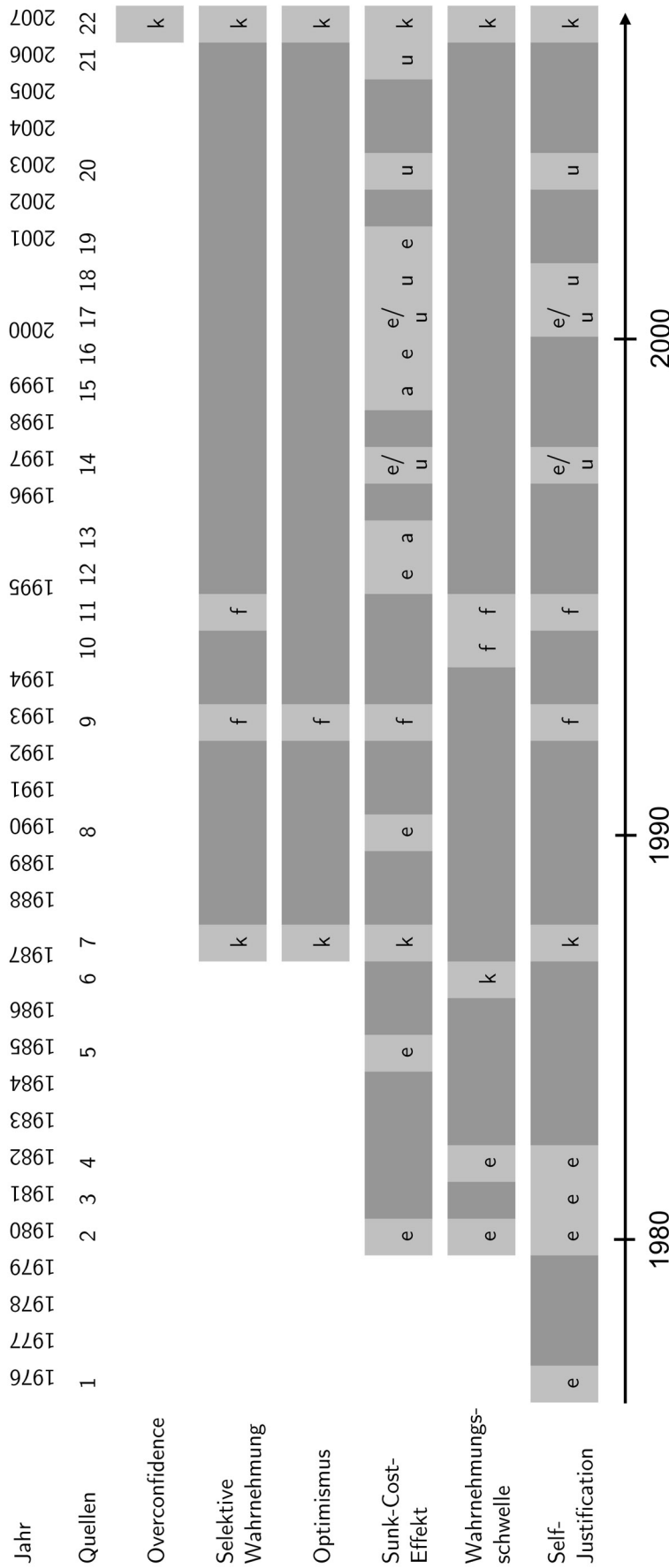
²⁷ Vgl. Lord 1989, S. 124.

²⁸ Vgl. Abschnitt 1.1 auf Seite 1.

²⁹ Kiesler/Sakumura 1966.

³⁰ Zur Definition des Begriffs „kognitiv“ vgl. Seite 3.

³¹ Die Ergebnisse der umfangreichen Forschung zu kognitiven Faktoren und Gegenmaßnahmen werden ausführlich in den Abschnitten 2.3.2.1 bis 2.3.4.2 auf den Seiten 30 bis 40 vorgestellt.



Legende: a = Archivdaten, e = Experiment, f = Fallstudie, k = konzeptionell, u = Umfrage

- 1 Staw 1976, S. 41 f. 2 Teger 1980, S. 16 und 52 ff. 3 Brockner/Rubin/Lang 1981, S. 77 f. 4 Brockner et alii 1982, S. 262 ff.
- 5 Arkes/Blumer 1985, S. 139. 6 Bowen 1987, S. 54 ff. 7 Staw/Ross 1987a, S. 48 ff. 8 Garland 1990, S. 729 ff. 9 Ross/Staw 1993, S. 715 ff.
- 10 Drummond 1995, S. 279. 11 Keil 1995, S. 431 ff. 12 Heath 1995, S. 53. 13 Staw/Hoang 1995, S. 486 ff. 14 Hollar 1997, S. 95
- 15 Camerer/Weber 1999, S. 60 ff. 16 Keil et alii 2000, S. 113 ff. 17 Hollar et alii 2000, S. 14 f. 18 Keil/Rai/Mann 2000, S. 650 ff.
- 19 Moon 2001, S. 110 f. 20 Keil et alii 2003, S. 259. 21 Dilts/Pence 2006b, S. 390 f. 22 Zayer 2007, S. 225.

Die Abbildung erhebt keinen Anspruch auf Vollständigkeit. Bei der Interpretation ist zu beachten, dass die Abbildung nicht die jeweils ersten Nennungen der kognitiven Effekte in der Literatur aufführt. Vielmehr sind solche Quellen dargestellt, in denen die genannten kognitiven Faktoren *im Zusammenhang mit dem Phänomen der Eskalation des Commitments* genannt werden. Unabhängig vom Konzept der Eskalation des Commitments wurden die kognitiven Effekte schon früher in der Literatur beschrieben.

Abbildung 2.1: Ausgewählte Publikationen zur Eskalation des Commitments

Nachdem diese Faktoren lange Zeit ungeordnet und teilweise widersprüchlich nebeneinander standen, klassifiziert Zayer erstmals in einem umfassenden, sachanalytisch hergeleiteten Modell die Ursachen für Eskalationstendenzen bei Projektfortführungen.³² Darüber hinaus erarbeitet Zayer theoriegeleitet Maßnahmen zur Verringerung der Eskalationstendenzen.³³ Die Maßnahmen haben das Ziel, die Prozessrationalität zu verbessern, und können dem Controlling im Sinne der Rationalitätssicherung zugeordnet werden.³⁴

In der aktuellen Managementliteratur wird die Eskalation des Commitments meist als Oberbegriff für das ökonomisch-irrationale Festhalten an scheiternden *Projekten* verstanden. So definieren beispielsweise Keil et alii Commitment dann als eskalierend, wenn Projekte weitergeführt werden, obwohl negative Informationen über die bisherigen Ressourcenallokationen vorliegen, die mit Unsicherheit in Bezug auf die Wahrscheinlichkeit der Zielerreichung gekoppelt sind.³⁵

Die vorliegende Arbeit folgt dem managementorientierten Fokus darin, die Eskalation des Commitments im Kontext von *Projekten* zu betrachten.³⁶ Gemäß der Norm DIN 69901 ist ein Projekt durch die Einmaligkeit der Bedingungen gekennzeichnet.³⁷ Diese Auffassung hat sich auch in der deutschsprachigen Betriebswirtschaftslehre etabliert.³⁸ Sehr ähnlich ist auch das Verständnis im englischsprachigen Raum, wo ein Projekt als befristete Anstrengung betrachtet wird, um ein einzigartiges Produkt, einen Service oder ein sonstiges Resultat zu erstellen.³⁹

In der Managementliteratur eng mit dem Begriff der Eskalation des Commitments verknüpft ist der Begriff der Projekteskalation. Allerdings werden hierbei die kognitiven Faktoren, die den Ursprung des Begriffs „Escalation of Commitment“ bilden, meist als *eine von mehreren* Arten von Einflussfaktoren auf die Projekteskalation angesehen. Weite Verbreitung hat beispielsweise die Taxonomie von Staw/Ross erlangt, die als Einflussfaktoren auf die Projekteskalation

³² Vgl. Zayer 2007, S. 151.

³³ Vgl. Zayer 2007, S. 213.

³⁴ Vgl. Weber/Schäffer 2006, S. 40.

³⁵ Vgl. Keil et alii 1995, S. 80 mit Referenz zu Brockner 1992. Ähnlich auch Schmidt/Calantone, die eskalierendes Commitment vorliegen sehen, wenn trotz Informationen, die einen erfolgreichen Abschluss als unwahrscheinlich erscheinen lassen, an einer scheiternden Vorgehensweise festgehalten wird; vgl. Schmidt/Calantone 2002, S. 104, mit Bezug auf Barton/Duchon/Dunegan 1989, Staw 1997 und Staw/Ross 1987a.

³⁶ Vgl. beispielsweise Cheng et alii 2003, Staw/Ross 1987a, Keil 1995 und Newman/Sabherwal 1996.

³⁷ Vgl. Norm DIN 69901: 1987-08. Für Beispiele von Aufgaben, die üblicherweise als Projekt bearbeitet werden, sowie für typische Merkmale von Projekten vgl. Demleitner 2006, S. 2 ff.

³⁸ Vgl. Hauschildt 2004, S. 84.

³⁹ Vgl. Project Management Institute (Hrsg.) 2006, S. 5.

Projektfaktoren, kognitive Faktoren, soziale Faktoren und organisationale Faktoren unterscheiden.⁴⁰ Die vorliegende Arbeit fokussiert sich auf den Erklärungsgehalt der *kognitiven* Faktoren für das ökonomisch-irrationale Festhalten an Projekten.

Für die Diskussion von ökonomisch-irrationalem Verhalten und zur Bestimmung eines rationalen Zeitpunkts für einen Projektabbruch stellt sich die Frage nach der Bedeutung des Begriffes „*rational*“. Obwohl der Begriff der Rationalität in vielen Disziplinen von zentraler Bedeutung ist, wie beispielsweise in der Philosophie und der Wissenschaftstheorie, hat sich bislang kein einheitliches Verständnis herausgebildet.⁴¹

McFadden beschreibt rationales Verhalten in den Wirtschaftswissenschaften recht allgemein als *vernünftig, geplant und konsistent*.⁴² In Anlehnung an die Betrachtung der Vernunft als Denken in Zweck-Mittel-Relationen von Gutenberg schlagen Weber/Schäffer eine konkretere Definition von „*Rationalität als herrschende Meinung von Fachleuten hinsichtlich einer bestimmten Zweck-Mittel-Situation*“⁴³ vor. Analog kann auch in Eskalationssituationen das Urteil unabhängiger Fachleute als Maßstab für rationales Entscheiden gelten.⁴⁴ Diesem Verständnis von Rationalität folgt die vorliegende Arbeit.

Kann ein Projekt entsprechend dem vorgestellten Rationalitätsverständnis als eskalierend identifiziert werden, sollte es abgebrochen werden. Ein *Projektabbruch* wird dabei als die Entscheidung definiert, die Arbeit an einem Projekt aufzugeben, bevor die ursprünglichen Ziele (vollständig) erreicht werden.⁴⁵

2.1.3 Verknüpfung der psychologischen Erklärungen mit der ökonomischen Theorie

Bei der Unterstützung für eine Entscheidung, wie beispielsweise über einen Projektabbruch, stand in der Vergangenheit in der Betriebswirtschaftslehre und insbesondere auch im Controlling häufig die Frage im Mittelpunkt, mit welchen Entscheidungsinstrumenten und -modellen

⁴⁰ Vgl. Staw/Ross 1987a, S. 67, Keil 1995, S. 422, und Newman/Sabherwal 1996, S. 27.

⁴¹ Vgl. hierzu und für eine ausführlichere Diskussion des Rationalitätsbegriffs Weber/Schäffer 2006, S. 42 ff.

⁴² Vgl. McFadden 1999, S. 73.

⁴³ Weber/Schäffer 2006, S. 43; Hervorhebung im Original.

⁴⁴ Vgl. Keil et alii 2003, S. 252.

⁴⁵ Vgl. Dean 1968, S. 118, und Tadisina 1986, S. 97.

eine rationale Bewertung von Alternativen und darauf basierend eine rationale Entscheidung sichergestellt werden kann. Exemplarisch dafür stellt beispielsweise Hubner die Frage in den Mittelpunkt: „Wie kann über die Weiterführung bzw. den Abbruch laufender F&E-Projekte durch die Formulierung und Lösung von Entscheidungsmodellen im Rahmen einer plandeterminierten Unternehmensführung entschieden werden?“⁴⁶ Zur Lösung dieses Problems wurde in Form von Investitionsrechnungen, Optionspreisen, dynamischer Programmierung und zahlreichen weiteren Ansätzen eine Vielzahl von mono- und multikriteriellen Entscheidungsinstrumenten entwickelt.⁴⁷

Im Gegensatz zur traditionellen ökonomischen Herangehensweise stellt sich aus verhaltensorientierter Perspektive die Frage, warum Individuen trotz aller entwickelter Entscheidungsinstrumente *systematisch nicht rational* über Projektabbrüche entscheiden. Eine mögliche Erklärung besteht darin, dass die gegenwärtig verbreiteten Controllinginstrumente und -systeme möglicherweise zu sehr der klassischen ökonomischen Sichtweise entsprechen und die kognitiven Schwächen menschlicher Entscheider noch nicht angemessen berücksichtigen.⁴⁸ Um Controllingsysteme den kognitiven Fähigkeiten der Akteure anzupassen, ist des demzufolge erforderlich, die klassischen ökonomischen Annahmen problembezogen zu modifizieren und die Konsequenzen für das Controlling abzuleiten.

Wie zuvor erläutert, ist dies auch der Anspruch, den die Forschung zum Behavioral Accounting erhebt: das klassische ökonomische Modell um verhaltenswissenschaftliche Komponenten zu erweitern.⁴⁹ Dies wird in der vorliegenden Arbeit, ebenso wie bei Zayer,⁵⁰ aufgegriffen. Im Folgenden wird dargelegt, wie diese Erweiterung des ökonomischen Modells mit den in Abbildung 2.1 auf Seite 13 dargestellten kognitiven Faktoren erfolgen kann.

Die Fundierung, welche Aspekte für die Eskalation des Commitments als relevant anzusehen sind, bietet dabei die psychologische Literatur. Wie in Abschnitt 2.1.2 auf Seite 12 dargestellt

⁴⁶ Hubner 1996, S. 23.

⁴⁷ Für eine Übersicht vgl. Hubner 1996, S. 154 ff. Für eine umfassende anwendungsorientierte Beschreibung von Entscheidungen mit (Software-)Tools vgl. beispielsweise Clemen/Reilly 2001. Für empirische Ergebnisse zum Einsatz von Investitionsbewertungsverfahren in der Praxis vgl. Weber et alii 2006, S. 43.

⁴⁸ Unerwünschte Reaktionen auf Controllingdaten werden häufig als dysfunktional oder opportunistisch betrachtet. Das entsprechende Verhalten wird als Problem angesehen, nicht das System. Dem hält Birnberg entgegen, dass unter Berücksichtigung des Umfelds, in dem der Manager handeln muss, seine Aktionen die angemessene Antwort auf unangemessene Controllingsysteme sein können; vgl. für diese Aussage in Bezug auf Accountingssysteme Birnberg 2000, S. 713.

⁴⁹ Vgl. zur geschichtlichen Entwicklung der Behavioral-Accounting-Forschung sowie zum klassischen ökonomischen Modell den Abschnitt 2.1.1 auf Seite 9.

⁵⁰ Vgl. Zayer 2007, S. 93.

Tabelle 2.1: Erweiterungen des ökonomischen Modells; in Anlehnung an Zayer 2007, S. 93

Annahme	Klassische Annahme der ökonomischen Theorie	Erweiterung	Annahme im erweiterten Modell
Resourceful	Der Akteur findet die relevanten Entscheidungsmöglichkeiten.	(1) Wahrnehmungsschwellen (2) selektive Wahrnehmung	Das Entscheidungsproblem ist nicht vollständig bekannt.
Restricted	Der Akteur hat Einkommensbeschränkungen.	keine Erweiterung	Übernahme der klassischen Annahme
Evaluating	Der Akteur kann frühere, jetzige und zukünftige Zustände in seinem Nutzenkalkül bewerten.	(3) Self-Justification (4) Sunk-Cost-Effekt	Je nach Frame der Situation ändert sich das Zielsystem des Entscheiders und der Selbstschutz tritt potentiell an die Stelle der Profitmaximierung. Bei der Bewertung können Fehler auftreten, da Sunk Costs zur Bewertung herangezogen werden.
Expecting	Der Akteur erwartet zukünftige Zustände und Geschehnisse auf der Basis bestimmter Wahrscheinlichkeiten.	(5) (übermäßiger) Optimismus (6) Overconfidence	Die Erwartungen werden durch übermäßigen Optimismus und Overconfidence verzerrt.
Maximizing	Der Akteur wählt die für ihn vorteilhafteste Entscheidung.	keine Erweiterung	Übernahme der klassischen Annahme

wurde, sind in den letzten Jahrzehnten in der psychologischen Forschung zahlreiche Schwächen im menschlichen Informationsverarbeitungsprozess identifiziert worden, die zu fehlerhaften Entscheidungen bei Investitionsprojekten führen und damit zur Projekteskalation beitragen.⁵¹ Während die kognitiven Schwächen in der psychologischen Literatur weitgehend fragmentiert und ohne übergeordneten Analyserahmen aufgeführt werden, verwendet Zayer bei der Entwicklung seines Modells die *Methode der abnehmenden Abstraktion*.⁵² Diese ermöglicht eine Verknüpfung zwischen der ökonomischen und der psychologischen Forschung. Dabei wird von Zayer das klassische ökonomische Modell (RREEMM/homo oeconomicus)⁵³ als Kerntheorie verwendet und dieses um problemspezifische Annahmen erweitert.⁵⁴ Tabelle 2.1 stellt die Erweiterungen des klassischen ökonomischen Modells dar.

⁵¹ Für einen ausführlichen Überblick vgl. Zayer 2007.

⁵² Vgl. zur Methode der abnehmenden Abstraktion Lindenberg (1991).

⁵³ Das Acronym RREEMM steht für resourceful, restricted, expecting, evaluating, maximizing man; vgl. beispielsweise Lindenberg 1985, Lindenberg 1990, S. 739, und Meckling 1976. Eine Erläuterung für die Adjektive bietet Tabelle 2.1.

⁵⁴ Für eine ausführliche Erläuterung der Ausdifferenzierung des homo oeconomicus mit begrenzten kognitiven Fähigkeiten vgl. Heine et alii 2006.

Die Verknüpfung psychologischer Erkenntnisse mit der ökonomischen Theorie bietet dabei nicht nur für die Ökonomie die Chance, realistischere Modelle zu entwickeln. Auch die Psychologie kann davon profitieren, wenn ihre separat stehenden Konzepte in eine übergeordnete Theorie eingebettet werden. So stellt Goethals fest: "Psychology, among other sciences, desperately needs unifying ideas, concepts that can usefully tie disparate literatures together."⁵⁵

Die einzelnen Erweiterungen werden in Abschnitt 2.3 näher erläutert. Basierend auf den empirischen Ergebnissen und statistischen Analysen erfolgt in Abschnitt 4.4.3 eine Beurteilung der Angemessenheit der hier vorgenommenen Erweiterungen des klassischen ökonomischen Modells.

2.1.4 Abgrenzung von nahe stehenden Arbeiten

In diesem Abschnitt werden kurz jene Arbeiten vorgestellt, die inhaltlich und forschungsmethodisch der vorliegenden Arbeit nahe stehen. Dies sind insbesondere die (wenigen) Umfragen, die zur Eskalation des Commitments durchgeführt wurden, sowie einige Experimente, die mit Strukturgleichungsmodellen ausgewertet wurden. Damit werden zwei Ziele verfolgt: Zum einen können so Überschneidungen ausgeschlossen werden und zum anderen kann der aktuelle Erkenntnisstand dieser ansatzweise vergleichbaren Forschungsprojekte illustriert werden.

Ein Experiment, dessen Ergebnisse mit einem Strukturgleichungsmodell analysiert wurden, stammt von Keil et alii.⁵⁶ In diesem Experiment wird der kulturelle Einfluss auf die Bereitschaft, ein Projekt fortzuführen, gemessen. Die zwei zentralen Ergebnisse lauten: (1) Die Wirkung von Risikoneigung auf die Risikowahrnehmung unterscheidet sich je nach Kultur in der Stärke. (2) Der Sunk-Cost-Effekt hingegen tritt in allen untersuchten Kulturkreisen ähnlich stark auf.

Methodisch zwischen Experiment und Umfrage liegen die Dissertation von Hollar und der darauf aufbauende Artikel von Hollar et alii.⁵⁷ Für seine Datenerhebung verwendet Hollar einen Fragebogen, in dem die Respondenten gebeten werden, hypothetische Situationen zu bewerten. Eine Aussage lautet beispielsweise: "A community college dean advocates a new vocational program due to high projected industry needs despite the program failing at other community

⁵⁵ Goethals 1992, S. 328.

⁵⁶ Vgl. Keil et alii 2000, S. 315.

⁵⁷ Vgl. Hollar 1997 und Hollar et alii 2000.

colleges”, wobei die Antworten von „1 = I very much disagree“ bis „6 = I very much agree“ skaliert sind. Aufbauend auf diesen Messungen analysieren Hollar et alii ein umfangreiches Strukturmodell mit verschiedenen Einflussfaktoren der Eskalation. Dabei zeigt sich, dass das Modell dann einen besonders guten Fit aufweist, wenn die psychologischen Faktoren als Mediatoren modelliert werden. Organisationale und soziale Faktoren wirken demnach mittelbar über psychologische Faktoren auf das Eskalationsverhalten.⁵⁸

Die Artikel von Keil/Rai/Mann und Keil et alii basieren auf einer Umfrage zur Eskalation von Softwareprojekten.⁵⁹ Als Einflussfaktoren auf die Eskalation werden dabei die Projektplanung, die Projektspezifikation, die Projektschätzung und die Projektkontrolle betrachtet. Dabei werden unter anderem auch Maßnahmen abgefragt, die auf kognitive Faktoren Einfluss haben können.⁶⁰ In der Auswertung wird geprüft, inwiefern sich verschiedene Konstrukte eignen, um zwischen scheiternden und erfolgreichen Projekten zu differenzieren. Die Ergebnisse zeigen, dass das Projektmanagement-Konstrukt die deutlichste Vorhersage erlaubt, ob ein Projekt scheitern wird. Aber auch andere Konstrukte, die auf verschiedenen verhaltenswissenschaftlichen Theorien basieren, erweisen sich als geeignet, um das Scheitern zu prognostizieren.

Wallace/Keil/Rai entwickeln ein Strukturgleichungsmodell für (Software-)Projektrisiken und prüfen es empirisch mit einer Umfrage. Auf kognitive Faktoren gehen sie dabei nicht ausdrücklich ein; die Studie sei hier dennoch erwähnt, weil dort unter dem Stichwort Planungs- und Kontrollrisiken Maßnahmen berücksichtigt sind, die auch bei der Reduzierung kognitiver Eskalationsfaktoren Wirkung zeigen können.⁶¹ Dazu zählt beispielsweise das Definieren von Meilensteinen, das in der Literatur als Maßnahme zur Reduzierung der Wahrnehmungsschwelle vorgeschlagen wird.⁶² Mit dem gleichen Datensatz führen Wallace/Keil/Rai eine Clusteranalyse für Softwareprojekte durch, um deren Risikoprofile zu klassifizieren.⁶³ Basierend auf der gleichen Umfrage publizieren Wallace/Keil einen kurzen Artikel zur Wirkung der Risiken auf die Projektergebnisse.⁶⁴ Dabei ist festzuhalten, dass soziale und technische Faktoren signifikant das Projektmanagementrisiko beeinflussen. Insbesondere der Projektumfang wirkt sich dabei auf viele andere Risiko-Dimensionen aus.

⁵⁸ Vgl. Hollar 1997 und Hollar et alii 2000.

⁵⁹ Vgl. Keil/Rai/Mann 2000 und Keil et alii 2003.

⁶⁰ Vgl. Kapitel 2.2 auf Seite 21.

⁶¹ Vgl. Wallace/Keil/Rai 2004a, S. 304.

⁶² Vgl. Abschnitt 2.3.2.1 auf Seite 31.

⁶³ Vgl. Wallace/Keil/Rai 2004b, S. 115 ff.

⁶⁴ Vgl. Wallace/Keil 2004, S. 68 ff.

Tabelle 2.2: Großzahlige Untersuchungen in verwandten Themenfeldern

Autoren	Thema	Forschungsmethode und Datengrundlage	Ausgewählte Kernergebnisse
Keil et alii 2000	Kulturelle Unterschiede in Bezug auf die Projekt-Fortsetzungsbereitschaft (Willingness to Continue a Project)	Experimente mit 579 Auditoren aus Finnland, den Niederlanden und Singapur	Die Wirkung von Risikoneigung auf die Risikowahrnehmung tritt in verschiedenen Kulturen unterschiedlich stark auf. Der Sunk-Cost-Effekt tritt sowohl in Kulturen mit mittlerer als auch mit geringer Uncertainty Avoidance auf.
Hollar 1997 und Hollar et alii 2000	Modelle des Eskalationsverhaltens	Umfrage mit hypothetischen Situationen (case scenarios), d. h. quasi-experimentell mit 239 College-Administratoren	Psychologische Faktoren können die Wirkung von organisationalen und sozialen Faktoren auf das Eskalationsverhalten mediieren.
Keil/Rai/Mann 2000 und Keil et alii 2003	Unterscheidung eskalierter und nicht eskalierter Projekte	Umfrage mit 579 Auditoren	Konstrukte mit Elementen aus Projektmanagement, Self Justification, Prospect Theory, Agency Theory und Approach Avoidance Theory können verwendet werden, um das Scheitern von Projekten vorherzusagen.
Wallace/Keil 2004 und Wallace/Keil/Rai 2004b	Risiken bei Softwareprojekten	Umfrage mit 507 Software-Projektmanagern	Soziale und technische Faktoren beeinflussen das Projektmanagementrisiko. Der Projektumfang ist ein besonders wichtiger Treiber des Risikos, der sich auf viele andere Dimensionen auswirkt.
Dilts/Pence 2006b	Projekteskalation in Abhängigkeit von der Position	Umfrage mit 32 Projektmanagern und 23 Geschäftsführern	Projektmanager und Executives nehmen Projekterfolg und Scheitern unterschiedlich wahr.

Basierend auf einer Umfrage unter Projektmanagern und Geschäftsführern analysieren Dilts/Pence verschiedene Faktoren, die beim Abbruch eines Projekts eine Rolle spielen. Dabei betrachten sie insbesondere den Einfluss der Position des Entscheidungsträgers auf seine Einschätzung kritischer Projektfaktoren, den Sunk-Cost-Bias und den Umgang mit Informationen. Die Analysen deuten darauf hin, dass Projektmanager und Geschäftsführer unterschiedliche Auffassungen vom Scheitern oder Erfolg eines Projekts haben. Die Anzahl der Datensätze beträgt dabei 55, so dass die Ergebnisse statistisch nur sehr eingeschränkt belastbar sind, wie die Verfasser hervorheben.⁶⁵

Tabelle 2.2 gibt einen Überblick über die vorgestellten Untersuchungen. Zusammenfassend ist festzustellen, dass bislang keine großzahlige Umfrage durchgeführt wurde, in der die Wirkung *kognitiver* Faktoren auf die Eskalation realer Projekte detailliert untersucht wurde. Dementspre-

⁶⁵ Vgl. Dilts/Pence 2006b.

chend besteht ein wesentlicher Beitrag für den wissenschaftlichen Erkenntnisfortschritt in der Untersuchung, ob die bisherigen konzeptionellen Überlegungen, Fallstudien und experimentellen Forschungen einer großzahligen empirischen Prüfung im realen Unternehmenskontext standhalten.

Wie zuvor erläutert, stellt das von Zayer entwickelte Modell den derzeit umfassendsten Erklärungsansatz zur Integration verschiedener kognitiver Faktoren und ihrer Wirkung auf die Projekteskalation dar. Zayers Modell wird deshalb als Ausgangspunkt für das in dieser Arbeit empirisch zu prüfende Modell herangezogen.⁶⁶ Im folgenden Abschnitt wird das Modell von Zayer vorgestellt und die vorgenommenen Vereinfachungen und Modifikationen erläutert.

2.2 Herleitung des Modells

Die theoretisch fundierte Herleitung eines Modells, das die Zusammenhänge zwischen den betrachteten Variablen abbildet, ist eine elementare Voraussetzung für die konfirmatorische Kovarianzstrukturanalyse.⁶⁷ Nachdem der angestrebte Erkenntnisgewinn bereits in Abschnitt 1.2 auf Seite 2 dargelegt wurde, wird deshalb in den folgenden Abschnitten die theoretische Fundierung eines Erklärungsmodells für die Projekteskalation vorgestellt. Den Ausgangspunkt bildet dabei, wie in Abschnitt 2.1.2 auf Seite 12 begründet, das von Zayer entwickelte Modell.⁶⁸

⁶⁶ Zur Relevanz einer Theorie für die Erklärung empirischer Daten merkt Coase sehr pointiert an: “Without a theory they had nothing to pass on except a mass of descriptive material waiting for a theory or a fire.”; Coase 1984, S. 230.

⁶⁷ So warnt beispielsweise Diller: „Es scheint freilich [...] so, dass nicht wenige Autoren den Versuchungen der modelltheoretisch messerscharfen und auswertungsanalytisch stringenten Untersuchungsanlage von Kausalanalysen erliegen, ohne zu erkennen, dass schon ihr entwickeltes Modell fehl spezifiziert, dem Untersuchungsgegenstand nicht angemessen ist oder keinen echten Erkenntnisgewinn für die Wissenschaft erbringt.“; Diller 2006, S. 611.

⁶⁸ Vgl. Zayer 2007.

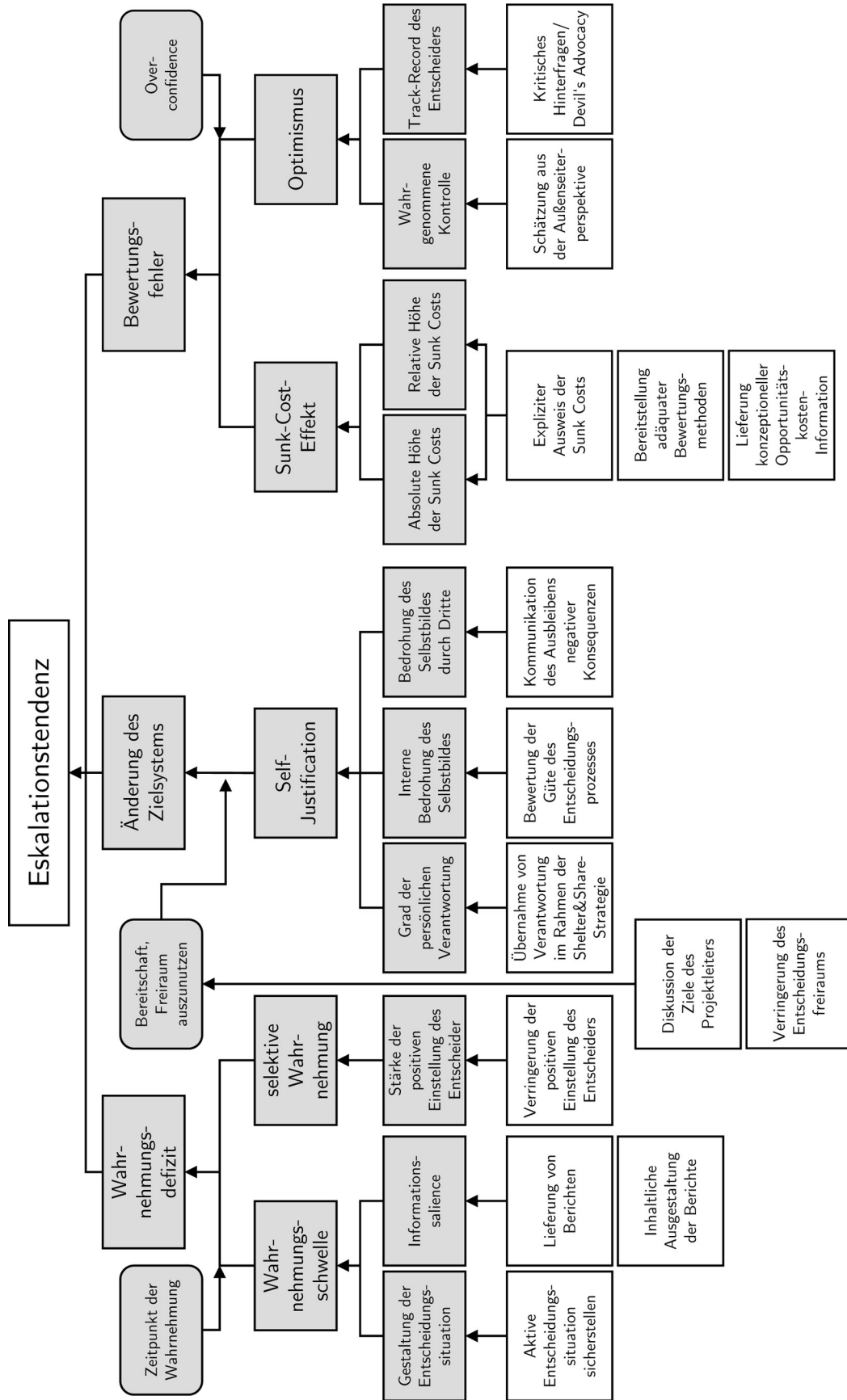


Abbildung 2.2: Gesamtmodell der Eskalationsfaktoren, Quelle: Zayer 2007, S. 225

2.2.1 Das Modell von Zayer als Ausgangspunkt

Abbildung 2.2 auf der vorherigen Seite zeigt das von Zayer entwickelte Gesamtmodell. Es besteht aus mehreren Ebenen. Die zentrale abhängige Variable ist die Eskalationstendenz. Darauf wirken verschiedene kognitive Faktoren, die hierarchisch in mehreren Detaillierungsebenen dargestellt werden. In der untersten Ebene sind Maßnahmen des verhaltensorientierten Controllings aufgelistet. Diese können dazu beitragen, das Auftreten des jeweiligen kognitiven Eskalationsfaktors zu vermeiden. Außerdem enthält das Modell drei moderierende Faktoren, die mit abgerundeten Ecken dargestellt sind: Zeitpunkt der Wahrnehmung, Bereitschaft, den Freiraum auszunutzen, und Overconfidence.

2.2.2 Modifikation des Modells für die empirische Prüfung

Um das Modell für die empirische Prüfung im Kontext risikoreicher Investitionsprojekte anzupassen, erscheint eine Reihe von Modifikationen angebracht, die im weiteren Verlauf der Arbeit vorgestellt und begründet werden. Insoweit davon einzelne Variablen oder Hypothesen betroffen sind, werden die Modifikationen in den entsprechenden Abschnitten in Kapitel 2.3 ab Seite 29 behandelt. Modifikationen, die sich auf das Gesamtmodell beziehen, werden im Folgenden vorgestellt.

2.2.2.1 Geltungsbereich des Modells

Der Fokus von Zayers Arbeit liegt auf unternehmensinternen Projekten der angewandten Forschung und Entwicklung.⁶⁹ Eine Erweiterung des Geltungsbereichs von Zayers Modell erscheint wünschenswert, denn aus wissenschaftlicher Sicht ist eine Theorie umso besser, je mehr Phänomene sie (bei gleicher Sparsamkeit der Annahmen) erklären kann.

Eskalation des Commitments tritt nicht nur in der Forschung und Entwicklung auf, sondern auch in zahlreichen anderen Kontexten.⁷⁰ Dies konstatiert auch Zayer⁷¹ und beschränkt sich

⁶⁹ Vgl. Zayer 2007, S. 10.

⁷⁰ Vgl. Abschnitt 1.1 auf Seite 1.

⁷¹ Zayer hält fest: „Zahlreiche Fallstudien aus anderen Bereichen verdeutlichen, dass verspätete Projektabbrüche stabile Verhaltensmuster sind, die auch unabhängig vom F&E-Kontext auftreten können. Vgl. z. B. die Aus-

deshalb für die Fundierung seines Modells nicht auf die F&E-Literatur, sondern fasst die Literaturrecherche weiter: „Dabei sollen auch dann nicht F&E-spezifische Studien betrachtet werden, solange vergleichbare Situationseigenschaften (wie beispielsweise hohe Unsicherheit oder eine geringe zur Verfügung stehende Datenbasis) vorliegen.“⁷²

Zayer nennt vier Charakteristika für die Abbruchentscheidung im F&E-Kontext: hohe inhärente Unsicherheit, hohe Bedeutung der Abbruchentscheidung, hohe Verantwortung des Projektleiters und große Entscheidungsfreiräume des Projektleiters.⁷³ Sind diese vier Merkmale stark ausgeprägt, so ist von einer hohen Erklärungskraft des Modells auszugehen. Demzufolge ist die F&E ein Gebiet, in dem das Deeskalationsmodell *besonders* aussagekräftig ist. Da die Merkmale jedoch auch in anderen Kontexten beobachtbar sind, ist davon auszugehen, dass das Modell für einen größeren Ausschnitt von Projekten Erklärungsgehalt aufweist.

Um die Übertragbarkeit des Modells von Zayer detaillierter beurteilen zu können, wurde folgende systematische Vorgehensweise angewendet: Mittels einer Volltextsuche wurde das gesamte Dokument von Zayer nach den Begriffen „F&E“ und „Forschung und Entwicklung“ durchsucht. Insgesamt wurden über 90 Aussagen zu F&E identifiziert. Für jede gefundene Textstelle wurde überprüft, welche F&E-spezifische Aussage für das Gesamtmodell getroffen wird. Die Aussagen wurden codiert, d. h., jede Aussage wurde einer übergeordneten Kategorie zugeordnet. Die Kategorien konnten auf insgesamt fünf Merkmale verdichtet werden, unter denen sich alle Einzelaussagen einordnen lassen. Diese fünf Merkmale beschreiben die Projekte, für die das Modell konzeptionelle Gültigkeit, d. h. inhaltliche Validität, aufweisen sollte. Demnach gilt das Modell für alle (1) mehrphasigen, (2) komplexen, (3) internen Investitionsprojekte (4) in von Unsicherheit geprägtem Umfeld, (5) bei denen der Entscheidungsverantwortliche (Mit-)Verantwortung für die Projektinitiierung und/oder die Durchführung bzw. Leistungserstellung trägt.

Diese Kriterien erfüllen nicht nur F&E-Projekte, sondern Projekte in einer Vielzahl von Kontexten. Selbst wenn nicht alle Kriterien erfüllt sind, können einzelne kognitive Faktoren einen relevanten Erklärungsgehalt aufweisen. Es erscheint plausibel anzunehmen, dass die meisten abgebrochenen bzw. gescheiterten Projekte mehrere dieser Kriterien erfüllen. Basierend auf dieser Annahme wird die vorliegende Untersuchung mit abgebrochenen und gescheiterten Projekten durchgeführt. Auf eine weitere Eingrenzung wurde bewusst verzichtet. Stattdessen wurden den

führungen bei Teger (1980), S. 99 ff., oder Berichte über verspätete Abbrüche in Belangen des alltäglichen Lebens bei Hall (1982), Pott (1992) oder Nutt (2002).“; Zayer 2007, S. 61.

⁷² Zayer 2007, S. 104.

⁷³ Zayer 2007, S. 26.

Teilnehmern zu Beginn des Fragebogens Beispiele für abgebrochene bzw. gescheiterte Projekte gegeben,⁷⁴ die den Situationen entsprechen, für die eine Eskalation des Commitments bereits nachgewiesen wurde.⁷⁵

Neben den kognitiven Eskalationsfaktoren umfasst das Modell von Zayer auch Controlling-Maßnahmen, die der Reduzierung dieser Faktoren dienen sollen. Auch sie sollen in der vorliegenden Studie auf ihre Wirksamkeit geprüft werden. Da bei sehr kleinen Unternehmen kein ausgeprägtes Controlling zu erwarten ist, bietet es sich an, die vorliegende Untersuchung auf mittlere und große Unternehmen zu fokussieren.⁷⁶ Die kognitiven Faktoren für sich genommen sind jedoch auch bei kleinen Unternehmen zu erwarten.

Zusammenfassend kann festgehalten werden: Das von Zayer entwickelte Modell bezieht sich zwar im Titel der Arbeit auf F&E, ist aber allgemeingültig konstruiert, so dass es für zahlreiche risikoreiche Projekte gelten kann. Keine der Modellkomponenten ist speziell auf F&E ausgerichtet und das beschriebene Phänomen der Eskalation des Commitments ist ein funktions- und branchenübergreifendes Problem.⁷⁷ Es ist allerdings davon auszugehen, dass die Eskalation des Commitments unter bestimmten Rahmenbedingungen stärker zum Tragen kommt als unter anderen. Projekte im Bereich der F&E stellen damit einen Spezialfall aus dem Geltungsbereich des Modells dar, für den dieses eine besonders hohe Erklärungskraft aufweist.

2.2.2.2 Modifikation des Entscheidungsprozessmodells

Das Modell von Zayer basiert auf einigen vereinfachenden Annahmen. Dies ist im Zuge einer klaren, analytischen Zusammenführung aller berücksichtigten Effekte unabdingbar. Im Rahmen einer empirischen Erhebung realer, komplexer Investitionsprojekte ist jedoch nicht davon auszugehen, dass alle Annahmen zutreffen. So geht das Modell von Zayer von einem vereinfachten, linearen Entscheidungsprozess ohne Rückkopplungen aus.⁷⁸ Zayer verweist jedoch darauf, dass der idealtypische Ablauf zwar in der betriebswirtschaftlichen Literatur etabliert ist, sequentiellen Abläufen in der psychologischen Literatur jedoch nur eine geringe deskriptive Validität für reales Entscheidungsverhalten zugesprochen wird.⁷⁹ In realen Projekten wird der

⁷⁴ Für den Wortlaut der Beispiele vgl. Seite 223.

⁷⁵ Für die Fälle, in denen eine Eskalation des Commitments dokumentiert wurde, vgl. Abschnitt 1.1 auf Seite 1.

⁷⁶ Vgl. Abschnitt 3.1.2 auf Seite 48.

⁷⁷ Vgl. Abschnitt 1.1 auf Seite 1.

⁷⁸ Vgl. Zayer 2007, S. 98.

⁷⁹ Vgl. Zayer 2007, S. 232, mit Verweis auf Simon 1981, S. 47 ff.

Entscheidungsprozess nicht immer der idealtypisch angenommenen Reihenfolge von Prozessschritten entsprechen. Vielmehr ist von Rückkopplungen auszugehen, die laut Zayer in einer nachfolgenden Arbeit näher untersucht werden müssen.⁸⁰

Darüber hinaus ist anzunehmen, dass nicht nur *ein* Durchlauf eines Entscheidungsprozesses stattfindet, sondern mehrfach über Fortgang oder Abbruch entschieden wird. Bei einer Aufhebung der genannten vereinfachenden Annahmen ist davon auszugehen, dass eine Wirkung der im Prozessmodell nachgelagerten Faktoren auf die vorgelagerten Faktoren nicht ausgeschlossen werden kann. Beispielsweise erscheint es plausibel, anzunehmen, dass ein einmal entstandenes Bedürfnis, seine bisherigen Entscheidungen zu rechtfertigen (Self-Justification),⁸¹ sich bei zukünftigen Durchläufen des Entscheidungsprozesses auf die Wahrnehmung der weiteren Feedbackinformationen auswirkt. Deshalb werden im Modell für die empirische Studie Kovarianzen zwischen den Faktoren ausdrücklich zugelassen.⁸² Welche möglichen gerichteten Wirkungen zwischen den Effekten einen guten Fit mit den empirischen Daten aufweisen, wird in Abschnitt 4.4.1 auf Seite 148 exploratorisch näher betrachtet.

Zayers Modell ist als entscheidungsbaumartige Struktur modelliert, bei der die einzelnen Effekte dichotom vorliegen, also entweder auftreten oder nicht auftreten.⁸³ Für reale Projekte erscheint es plausibler, davon auszugehen, dass die kognitiven Faktoren kontinuierliche Phänomene sind, also in verschiedenen starken Ausprägungen auftreten können. Die Variablen werden deshalb nicht dichotom operationalisiert, sondern als quasi-kontinuierliche Skalen mit sieben Antwortmöglichkeiten, die verschieden starke Ausprägungen messen.⁸⁴

2.2.2.3 Reduzierung der Elemente

Zayer gliedert sein Modell in vier Ebenen unterhalb der Eskalation (vgl. Abbildung 2.2 auf Seite 22).⁸⁵ Die Ebenen bezeichnet er dabei als Prozessschritt, psychologischer Effekt, Faktor und Maßnahmen.⁸⁶ Für die empirische Untersuchung im Rahmen dieser Studie muss die Anzahl der Elemente reduziert werden. Dies hat vor allem zwei Gründe: Zum einen wird die Anzahl

⁸⁰ Vgl. Zayer 2007, S. 98 f.

⁸¹ Vgl. hierzu die Erläuterungen zu Self-Justification in Abschnitt 2.3.3.1 auf Seite 35.

⁸² Vgl. Abschnitt 4.4.1.2 auf Seite 150.

⁸³ Vgl. Zayer 2007, S. 98 und die Abbildung bei Zayer 2007, S. 99.

⁸⁴ Vgl. Abschnitt 3.2.2.5 auf Seite 59.

⁸⁵ Vgl. Zayer 2007, S. 225.

⁸⁶ Vgl. Zayer 2007, S. 154 f.

der Fragen durch die zumutbare Länge des Fragebogens begrenzt,⁸⁷ zum anderen liegen einige Elemente des Modells konzeptionell sehr nah beieinander (beispielsweise Wahrnehmungsdefizit und Wahrnehmungsschwelle). Dies kann im Extremfall zu quasi-tautologischen Hypothesen über die Beziehung zwischen den Variablen führen.⁸⁸ Darüber hinaus bewirken eng verwandte Konstrukte häufig Irritationen bei den Befragten, wenn die feinen Nuancen nicht nachempfunden werden können und die Respondenten sich darüber wundern oder ärgern, mehrmals das Gleiche gefragt zu werden.⁸⁹ Um dieser Gefahr bereits in einem frühen Stadium zu begegnen, wurde ein Pretest zur Ermittlung der Trennschärfe zwischen den vorgesehenen Konstrukten durchgeführt (siehe Abschnitt 3.4.1.2 auf Seite 79).

Homburg/Klarmann weisen außerdem darauf hin, dass die Variablen auch deshalb nicht zu ähnlich konzeptionalisiert werden sollten, da sonst Multikollinearität, d. h. starke Korrelationen zwischen den jeweiligen Variablen, auftreten können, die zu ungenauen und instabilen Ergebnissen führen.⁹⁰

Da sowohl die Bearbeitungsdauer des Fragebogens als auch die konzeptionelle Dimension eine Reduzierung der Elemente nahe legt, wurde das Modell von Zayer vereinfacht. Dabei wurde darauf Wert gelegt, dass die Breite des Erklärungsansatzes erhalten bleibt. Eine Fokussierung auf beispielsweise einen einzelnen kognitiven Faktor erschien deshalb nicht als angebracht. Stattdessen wurde die Entscheidung getroffen, aus dem Modell von Zayer⁹¹ die zweite und vierte Ebene (von oben nach unten gezählt) zu entfernen.⁹²

2.2.3 Das theoretische Modell für die empirische Prüfung

Das Modell für die empirische Prüfung als Resultat der zuvor diskutierten Vereinfachungen und leichten Modifikationen ist in Abbildung 2.3 auf der nächsten Seite dargestellt. Das Modell zielt

⁸⁷ Vgl. hierzu den Pretest zur Ausfülldauer in Abschnitt 3.4.1.5 auf Seite 89.

⁸⁸ Vgl. Homburg/Klarmann 2006, S. 728.

⁸⁹ Vgl. Diller 2006, S. 615.

⁹⁰ Vgl. Homburg/Klarmann 2006, S. 729. Im Kapitel zur Datenanalyse wird die Problematik der Trennschärfe noch einmal aufgegriffen. Die Unterschiedlichkeit der gemessenen Konstrukte wird dort mit Hilfe des Fornell/Larcker-Kriteriums analysiert; vgl. Abschnitt 4.3.5.2 auf Seite 146.

⁹¹ Vgl. Abbildung 2.2 auf Seite 22.

⁹² Die vierte Ebene wurde von Beginn an ausgeschlossen, die zweite Ebene nach dem Sortierpretest; vgl. Abschnitt 3.4.1.2 auf Seite 79.

auf die Erklärung der abhängigen Variable Projekteskalation ab. Des Weiteren umfasst das Modell die fünf kognitiven Eskalationsfaktoren Wahrnehmungsschwelle, selektive Wahrnehmung, Self-Justification-Druck, Sunk-Cost-Effekt und Optimismus. Die vorliegende Studie beschränkt sich bei der Betrachtung dieser kognitiven Faktoren, ebenso wie Zayer 2007, auf einzelne Entscheidungsträger. Gruppeneffekte werden zur Komplexitätsreduzierung nicht berücksichtigt.

Jedem kognitiven Faktor ist eine Reihe von Maßnahmen zugeordnet, die in der Unternehmenspraxis helfen sollen, die Stärke des Auftretens des Faktors zu verringern. Schließlich beinhaltet das Modell auch drei moderierende Faktoren: Zeitpunkt der Wahrnehmung, nicht hinterfragter Entscheidungsfreiraum und Overconfidence.

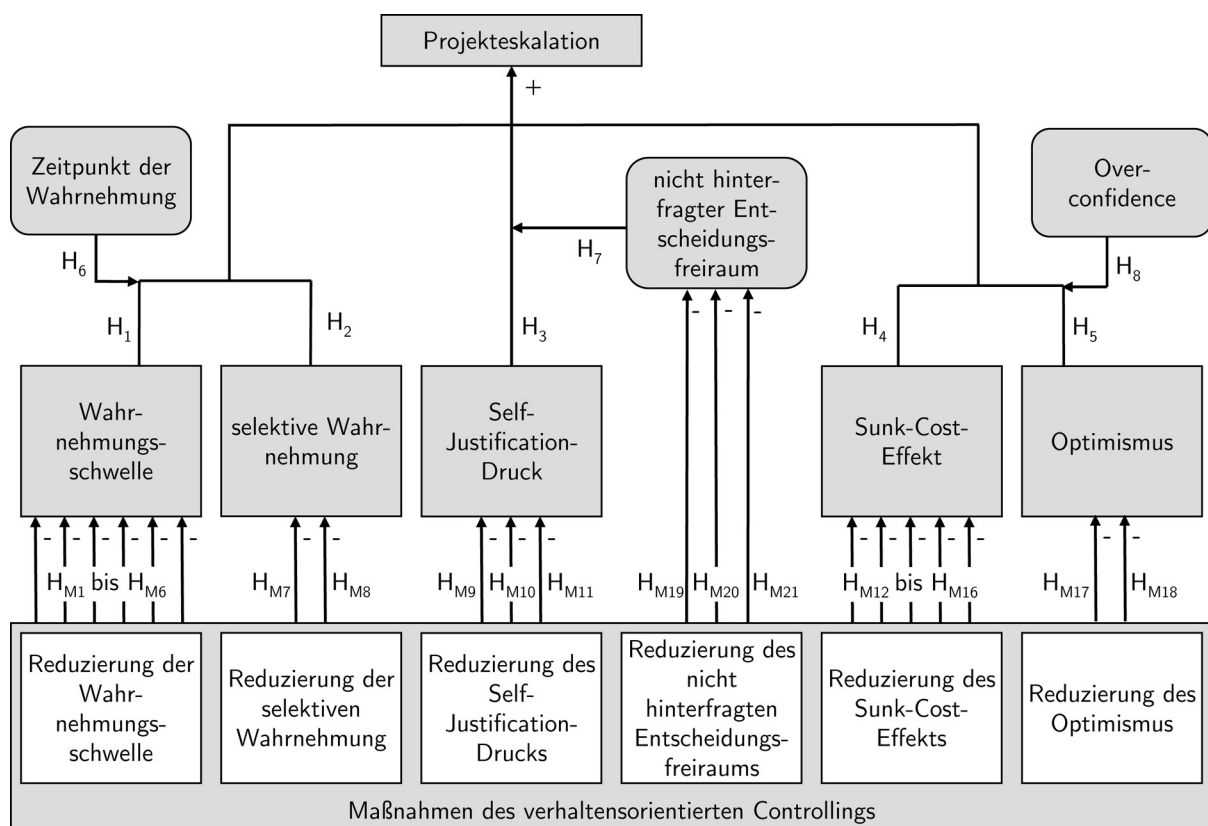


Abbildung 2.3: Eskalationsmodell für die empirische Untersuchung; in Anlehnung an Zayer 2007, S. 213 (vereinfacht und leicht modifiziert)

Die in Abbildung 2.3 dargestellten Variablen und Hypothesen werden in den folgenden Abschnitten erläutert. Die beiden Variablen Self-Justification-Druck und nicht hinterfragter Entscheidungsfreiraum tragen in dem Modell für die empirische Prüfung eine andere Bezeichnung als in Zayers Originalmodell; die Gründe dafür sind in den Abschnitten 2.3.3.1 und 2.3.5.2 dargelegt.

2.3 Erläuterung der Variablen und Herleitung der Hypothesen

Um die vorliegende Arbeit nachvollziehen zu können, ist ein grundlegendes Verständnis der untersuchten Zusammenhänge unabdingbar. Die folgenden Abschnitte greifen deshalb die von Zayer dargelegten Wirkungsmechanismen auf, verweisen auf einen Teil der von Zayer herangezogenen Quellen und integrieren weitere unterstützende Literatur. Zusätzlich werden die von Zayer graphisch dargestellten Zusammenhänge in explizite Hypothesen überführt, die im späteren Verlauf der Arbeit empirisch geprüft werden. Um den umfangbezogenen Rahmen nicht zu sprengen, können die Variablen und die Herleitung der Hypothesen nur knapp skizziert werden,⁹³ für die vollständige Darstellung sei auf Zayer 2007 verwiesen.⁹⁴

2.3.1 Eskalationstendenz und Projekteskalation

Die Variable *Projekteskalation* bezieht sich dabei auf das Phänomen der Eskalation des Commitments⁹⁵ und dient als Oberbegriff für die „Tendenz zu irrational verspäteten Projektabbrüchen“.⁹⁶ Ein Projekt gilt dabei als verspätet abgebrochen, wenn es fortgeführt wird, obwohl eine rationale unternehmerische Bewertung durch Fachleute einen Abbruch nahelegen würde.⁹⁷ Dies ist beispielsweise der Fall, wenn die Bewertung mit Real-Optionen, Scoring-Verfahren oder der Kapitalwertmethode zu dem Ergebnis kommt, dass ein Abbruch vorteilhaft ist.⁹⁸

Mit der Verwendung der Projekteskalation als abhängiger Variable hängt zusammen, dass die gemessene Eskalation nicht nur die kognitiven Wirkungen umfasst, sondern auch andere Faktoren als Ursache zulässt.⁹⁹ Dies wiederum bedeutet, dass das für die empirische Erhebung verwendete Modell *nicht* den Anspruch erhebt, die gemessene Eskalation vollständig zu erklären. Vielmehr stellen die empirisch gemessenen kognitiven Faktoren im Modell *eine* Kate-

⁹³ Herleitung, Erläuterung und Diskussion des Modells umfassen bei Zayer 238 Seiten.

⁹⁴ Vgl. Zayer 2007.

⁹⁵ Für ausführlichere Erläuterungen zu den Begriffen Eskalation des Commitments und Projekteskalation vgl. Abschnitt 2.1.2 auf Seite 12.

⁹⁶ Zayer 2007, S. 64.

⁹⁷ Für die hier verwendete Definition von Rationalität vgl. Abschnitt 2.1.2 auf Seite 12.

⁹⁸ Vgl. Zayer 2007, S. 54.

⁹⁹ Vgl. hierzu auch die Erläuterungen zum Begriffsverständnis in Abschnitt 2.1.2 auf Seite 12.

gorie möglicher Einflussfaktoren dar. Wie hoch ihr Erklärungsanteil für die Projekteskalation ist, wird mit den statistischen Analysen zu klären sein.¹⁰⁰

Die abhängige Variable in Zayers Modell trägt die Bezeichnung *Eskalationstendenz*. Da dieser Begriff nicht ausdrückt, auf welche Analyseeinheit er sich bezieht (einzelner Entscheidungsträger, Projekt, Unternehmen etc.), wird stattdessen für die vorliegende Arbeit der Begriff Projekteskalation verwendet. Die Analyseeinheit für die Projekteskalation stellt das einzelne *Projekt* dar.¹⁰¹ Im Gegensatz dazu werden die im Folgenden erläuterten kognitiven Faktoren auf der Ebene des *einzelnen Verantwortlichen* gemessen.

2.3.2 Wahrnehmungsdefizite

Eine korrekte Bewertung setzt voraus, dass der Entscheidungsverantwortliche alle Aspekte des (meist mehrdeutigen) Feedbacks wahrnimmt. Im menschlichen Wahrnehmungsprozess kommen jedoch verschiedene Filter zum Tragen, die zu Wahrnehmungsdefiziten führen können.¹⁰² Insbesondere wenn das Feedback sehr umfangreich und mehrdeutig ist, können die kognitiven Beschränkungen von Individuen erhebliche Wahrnehmungsdefizite hervorrufen. Dies wiederum kann dazu führen, dass bei den Verantwortlichen kein Problembewusstsein entsteht und dementsprechend auch keine Maßnahmen ergriffen werden.¹⁰³ Zayer identifiziert zwei Ursachen für das Wahrnehmungsdefizit: die Existenz von Wahrnehmungsschwellen und die selektive Wahrnehmung.

2.3.2.1 Wahrnehmungsschwelle

Damit Menschen einen Reiz wahrnehmen können, muss dieser aus dem „Hintergrundrauschen“ anderer Informationen hervorstechen, also eine Wahrnehmungsschwelle überschreiten.¹⁰⁴ Um eine Wirkung hervorzurufen, muss die Information gleichzeitig so beschaffen sein, dass

¹⁰⁰ Vgl. Abschnitt 4.4.1 auf Seite 148.

¹⁰¹ Für Studien, die ebenfalls Projekte als Analyseeinheit haben, vgl. beispielsweise Keil/Robey 1999, S. 71, und Keil et alii 2003, S. 254.

¹⁰² Vgl. Jost 2000, S. 101 ff., und Zayer 2007, S. 105.

¹⁰³ Vgl. Montealegre/Keil 2000, S. 432.

¹⁰⁴ Vgl. Zimbardo/Gerig 1996, S. 118.

sie als verhaltensrelevant identifiziert werden kann.¹⁰⁵ Wenn die negativen Informationen über ein scheiterendes Projekt zu unauffällig sind, besteht somit die Gefahr, dass das Projekt nicht abgebrochen wird, weil der Entscheidungsverantwortliche die Probleme nicht wahrnimmt oder nicht als handlungsrelevant erkennt.¹⁰⁶ Je höher demzufolge die Wahrnehmungsschwelle ist, desto stärker ist die zu erwartende Projekteskalation.

Im Folgenden wird für die Hypothesen die Bezeichnung „positive Wirkung“ für Beziehungen verwendet, wenn die abhängige Variable umso größer ist, je größer die unabhängige Variable ist. Der Begriff „positiv“ stellt also keine Wertung im Sinne von „gut“ dar, sondern die Richtung der Wirkung.

H₁: Die Höhe der Wahrnehmungsschwelle wirkt direkt positiv auf die Projekteskalation.

Im Kontext von Projektabbrüchen können gemäß dem Modell von Zayer zwei Faktoren die Höhe der Wahrnehmungsschwelle beeinflussen: zum einen die Gestaltung der Entscheidungssituation, zum anderen die Informationssalienc.

Gestaltung der Entscheidungssituation

In Bezug auf die Entscheidungssituation kann zwischen einer passiven und einer aktiven Situation unterschieden werden.¹⁰⁷ Während bei einer passiven Entscheidungssituation das Projekt prinzipiell weiterläuft, wenn keine gegenteiligen Maßnahmen ergriffen werden, wird bei einer aktiven Situation eine bewusste Entscheidung erzwungen.¹⁰⁸ Letzteres kann helfen, die Eskalation zu reduzieren, indem die Entscheidungsverantwortlichen dazu veranlasst werden, Informationen zu beachten, die andernfalls unter der Wahrnehmungsschwelle geblieben wären.¹⁰⁹

¹⁰⁵ Singer 2002a erläutert dies wie folgt: „Wie jeder aus eigener Erfahrung weiß, wird nur ein kleiner Teil dessen, was jeweils wahrnehmbar ist, auch tatsächlich wahrgenommen, und wiederum nur ein kleiner Teil des Wahrgenommenen wird erinnert. Spuren im Gedächtnis hinterlassen nur die Ereignisse, auf die sich die Aufmerksamkeit gerichtet hatte und die im Verhaltenskontext als bedeutsam erkannt wurden.“; Singer 2002a, S. 168.

¹⁰⁶ Vgl. Zayer 2007, S. 106.

¹⁰⁷ Vgl. Zayer 2007, S. 107.

¹⁰⁸ Vgl. Brockner/Shaw/Rubin 1979.

¹⁰⁹ Vgl. Brockner/Shaw/Rubin 1979 und Rubin et alii 1980.

Zayer führt mehrere Maßnahmen an, die in der Literatur als wirkungsvoll zur Verbesserung der Entscheidungssituation identifiziert werden.¹¹⁰ Als Voraussetzung sollten zunächst Meilensteine für das Projekt geplant werden.¹¹¹ An den Meilensteinen sollten Genehmigungen für die Fortsetzung erforderlich sein oder auf andere Weise sichergestellt werden, dass aktive Entscheidungen über Fortsetzung oder Abbruch gefällt werden (stage/gate approval).¹¹² Damit soll verhindert werden, dass ein scheiterndes Projekt mehr oder weniger unbemerkt oder unhinterfragt weiterlaufen kann.

Aus den diskutierten Maßnahmen ergeben sich die Hypothesen H_{M1} bis H_{M3} . Der Begriff „negativ“ bezeichnet dabei die Wirkungsrichtung und ist nicht wertend zu interpretieren. Im Gegenteil, aus Sicht des Controllings ist eine signifikante negative Wirkung das Ziel, also „gut“. Hypothesen für Maßnahmen werden durch ein tiefgestelltes „M“ gekennzeichnet.

H_{M1} : Meilensteine wirken direkt negativ auf die Wahrnehmungsschwelle.

H_{M2} : Genehmigungen für die Fortführung des Projekts wirken direkt negativ auf die Wahrnehmungsschwelle.

H_{M3} : Aktive Entscheidungen über die Fortführung des Projekts wirken direkt negativ auf die Wahrnehmungsschwelle.

¹¹⁰ Für die Formulierung der Hypothesen ist im Hinblick auf die spätere empirische Prüfung eine methodische Grundregel zu beachten. Fragestellungen sollten für eine empirische Untersuchung stets eindeutig sein und nur genau einen Inhalt haben. Einige Maßnahmen, die Zayer vorschlägt, umfassen jedoch mehrere Facetten. So lautet beispielsweise Maßnahme 1 „Planung des Projekts entlang eines Meilenstein-Prozesses und Sicherstellen einer aktiven Entscheidung über die Fortführung an den Meilensteinen“; Zayer 2007, S. 177. Würde diese Maßnahme in dieser Formulierung abgefragt werden, könnte dies die Validität senken, da bei einigen Befragten möglicherweise die Planung mit Meilensteinen erfolgte, jedoch keine aktive Entscheidung stattfand. Für die Befragten ist die Antwort dann gegebenenfalls nicht eindeutig. Deshalb wurden in diesem und ähnlichen Fällen die Vorschläge von Zayer bereits während der Hypothesenformulierung in Einzelkomponenten zerlegt, die später separat abgefragt werden können.

¹¹¹ Vgl. Cooper/Edgett/Kleinschmidt 2002a und Cooper/Edgett/Kleinschmidt 2002b. Wallace/Keil/Rai konzeptionalisieren das unklare Definieren von Meilensteinen als ein Element des Planungs- und Kontrollrisikos; vgl. Wallace/Keil/Rai 2004a, S. 299.

¹¹² Vgl. McGrath/Ferrier/Mendelow 2004, S. 94. Keil et alii betrachten Meilensteine einmal im Zusammenhang mit „Project Planning“ und zum anderen unter der Bezeichnung „Regular Checkpoints“ als Element des Konstrukts „Project Monitoring“. In ihrer Umfrage mit 579 Auditoren zeigte das Konstrukt „Project Planning“ keinen Erklärungsgehalt für die Projekteskalation, das Konstrukt „Project Monitoring“ hingegen starken Erklärungsgehalt; vgl. Keil et alii 2003, S. 255 ff.

Steigerung der Informationssalience

Wichtige Informationen können umso eher die Wahrnehmungsschwelle überwinden, je auffälliger sie im Vergleich zu anderen Informationen sind. Der Grad der Auffälligkeit wird dabei als Saliencce bezeichnet.¹¹³ Schon 1970 betonten Jones/McLean die Notwendigkeit, nicht nur Ziele und Zeitpläne festzulegen, sondern diese auch für alle Beteiligten präsent zu halten.¹¹⁴ Um die Eskalation zu reduzieren, muss deshalb die Saliencce der Projektinformationen erhöht werden.¹¹⁵ In der konzeptionellen Literatur und der experimentellen Forschung können mehrere Maßnahmen ermittelt werden, die zu einer höheren Saliencce der entscheidungsrelevanten Informationen führen.¹¹⁶ Zum einen kann die regelmäßige Erstellung und Lieferung von Projektfortschrittsberichten die Eskalation reduzieren, vor allem wenn negative Informationen frühzeitig salient präsentiert werden.¹¹⁷ Dazu bietet es sich beispielsweise an, Kosten, Qualität, Leistung und Zeitaufwand aufzuführen.¹¹⁸

Hantula/Bragger weisen in einem Laborexperiment nach, dass mehrdeutiges Feedback zu stärkerer Eskalation führt.¹¹⁹ Deshalb ist es wichtig, die Aussagekraft der Projektfortschrittsberichte zu erhöhen und die Berichte so klar und eindeutig wie möglich zu gestalten.¹²⁰ Dies kann insbesondere durch die Lieferung eines eindeutigen Beurteilungskriteriums (Zielwert) und die Quantifizierung der Feedback-Information erfolgen.¹²¹ Keil/Robey bezeichnen diese Maßnahme als „Clarity of Criteria for Success and Failure“.¹²²

H_{M4}: Projektfortschrittsberichte wirken direkt negativ auf die Wahrnehmungsschwelle.

H_{M5}: Eindeutige Beurteilungskriterien für die Projektinformationen wirken direkt negativ auf die Wahrnehmungsschwelle.

H_{M6}: Die Quantifizierung der Informationen über den Projektverlauf wirkt direkt negativ auf die Wahrnehmungsschwelle.

¹¹³ Vgl. Zayer 2007, S. 108 f.

¹¹⁴ Vgl. Jones/McLean 1970, S. 8.

¹¹⁵ Vgl. Ghosh 1995, S. 53.

¹¹⁶ Vgl. Zayer 2007, S. 177 f.

¹¹⁷ Vgl. Brockner et alii 1982, Drummond 1995, Ghosh 1997, Rubin/Brockner 1975, Staw/Ross 1987a und Teger 1980.

¹¹⁸ Vgl. Riedl 1990, S. 116 ff.

¹¹⁹ Vgl. Hantula/Bragger 1999, S. 434 ff.

¹²⁰ Vgl. Bowen 1987, Ghosh 1997, S. 90, und Staw/Ross 1987a, S. 73.

¹²¹ Vgl. Ghosh 1997, Keil/Robey 1999, S. 74, und Kernan/Lord 1989, S. 1129.

¹²² Keil/Robey 1999, S. 69.

2.3.2.2 Selektive Wahrnehmung

Neben der Wahrnehmungsschwelle ist die selektive Wahrnehmung der zweite Faktor in Zayers Modell, der zu einem Wahrnehmungsdefizit beiträgt.¹²³ Dieser Bias führt dazu, Informationen zu suchen, die den bereits existierenden Standpunkt verstärken und gegenläufige Informationen zu vermeiden.¹²⁴ Im Englischen wird die selektive Wahrnehmung auch als Confirmation Bias oder Confirming Evidence Bias bezeichnet.¹²⁵

Wie stark Manager der selektiven Wahrnehmung unterliegen, ist umstritten.¹²⁶ Experimente mit Studenten zeigen, dass die selektive Wahrnehmung durch eine fachbezogene Ausbildung gesenkt werden kann.¹²⁷ Ausgehend von der Annahme, dass die Mehrzahl der Entscheidungsverantwortlichen eine positive Einstellung zu ihren Projekten hat, führt die selektive Wahrnehmung zur Vermeidung projektkritischer Einsichten und steigert damit die Gefahr der Projekteskalation.

H₂: Die selektive Wahrnehmung wirkt direkt positiv auf die Projekteskalation.

Zayer gibt als Ansatzpunkt zur Reduzierung der selektiven Wahrnehmung die Einstellung des Entscheiders an.¹²⁸

Veränderung der Einstellung des Empfängers

Da die eskalationssteigernde Wirkung der selektiven Wahrnehmung nur dann auftritt, wenn der Entscheidungsverantwortliche eine positive Einstellung zum Projekt hat, bietet es sich an, bei der Einstellung des Verantwortlichen anzusetzen.¹²⁹ Um eine realistische Einstellung zu

¹²³ Vgl. Zayer 2007, S. 111 ff. In den Neurowissenschaften wird der Begriff selektive *Aufmerksamkeit* verwendet, den Singer wie folgt erläutert: „Bestehen bestimmte Erwartungen, richtet sich die Aufmerksamkeit auf die Sinneskanäle, welche die erwarteten Ereignisse übertragen werden. Diese kommen dann bevorzugt, aber natürlich auf Kosten anderer Vorgänge zur Verarbeitung, die erwarteten Inhalte werden schneller verarbeitet, schneller identifiziert und gelangen dann meist auch bevorzugt ins Bewußtsein [sic] und in die Langzeitspeicher.“; Singer 2002b, S. 80.

¹²⁴ Vgl. Hammond/Keeney/Raiffa 2006, S. 123.

¹²⁵ Vgl. Camerer 1995 und Hammond/Keeney/Raiffa 2006, S. 123.

¹²⁶ Walsh argumentiert beispielsweise auf der Basis eigener Experimente, dass die von Dearborn/Simon 1958 unternommene Studie entgegen deren Auffassung *keinen* Beleg für eine selektive Wahrnehmung der Manager liefert; vgl. Walsh 1988, S. 889.

¹²⁷ Vgl. Cloyd/Spilker 2000.

¹²⁸ Vgl. Zayer 2007, S. 181 ff.

¹²⁹ Vgl. Bazerman 1993, S. 419 f., Brockner/Rubin 1985, S. 195, Kadous/Sedor 2004 und Nathanson et alii 1982.

erzielen, sollte insbesondere die Kommunikation der Wahrscheinlichkeit eines Fehlschlags erfolgen.¹³⁰ So zeigen beispielsweise Experimente, dass Personen, die über das Phänomen verspäteter Projektabbrüche informiert werden, Projekte früher abbrechen.¹³¹ Darüber hinaus brechen Personen, die explizit aufgefordert werden, nach abbruchrelevanten Kriterien zu suchen, ein Projekt früher ab als solche Personen, die die gleichen Informationen ohne diesen expliziten Hinweis analysieren.¹³²

H_{M7}: Hinweise auf die Gefahr der Projekteskalation wirken direkt negativ auf die selektive Wahrnehmung.

H_{M8}: Hinweise auf die Möglichkeit eines Projektabbruchs wirken direkt negativ auf die selektive Wahrnehmung.

2.3.3 Änderung des Zielsystems

Wenn das Wahrnehmungsdefizit vermieden werden konnte, stellt gemäß Zayer die Änderung des Zielsystems die zweite große Eskalationsgefahr dar.¹³³ Der Entscheidungsverantwortliche ist sich nun bewusst, dass es Schwierigkeiten mit dem Projekt gibt. Unter bestimmten Umständen, die im Folgenden näher erläutert werden, kann es jedoch vorkommen, dass eine Änderung des Zielsystems des Verantwortlichen auftritt.¹³⁴ Unter der Annahme, dass der Verantwortliche ursprünglich das Ziel hatte, im Interesse des Unternehmens zu handeln, kann eine Änderung seiner Ziele dahingehend erfolgen, dass er nach dem Bewusstwerden negativer Projektinformationen das Projekt aus persönlichen Gründen weiter führt, obwohl dies dem ökonomischen Interesse des Unternehmens widerspricht. Ursache dafür ist der Self-Justification-Effekt.

2.3.3.1 Self-Justification-Effekt

Der Self-Justification-Effekt bezieht sich auf das menschliche Bestreben, konsistent zu wirken und das Selbstbild zu schützen. Dies kann dazu führen, dass Projekte auch dann weitergeführt

¹³⁰ Vgl. Brockner/Rubin 1985, S. 193 ff.

¹³¹ Vgl. Nathanson et alii 1982.

¹³² Vgl. Kadous/Sedor 2004.

¹³³ Vgl. Zayer 2007, S. 115 ff.

¹³⁴ Vgl. Teger 1980, S. 15.

werden, wenn dies aus ökonomischer Sicht nicht zu rechtfertigen ist.¹³⁵ Durch zusätzliche Investitionen wird dabei der Eindruck erweckt bzw. verstärkt, die vorherigen Investitionsentscheidungen seien korrekt gewesen. Der Self-Justification-Effekt bildet den Ursprung der Eskalationsforschung¹³⁶ und geht zurück auf Arbeiten zum Commitment von Kiesler/Sakumura.¹³⁷ Diese wiederum bauen auf Festingers¹³⁸ Theorie der kognitiven Dissonanz auf.¹³⁹ Der Begriff vom Schutz des Selbstbilds bezieht sich hier sowohl auf das positive Bild, das der Verantwortliche von sich selbst hat, als auch auf den Eindruck, den andere von ihm haben.¹⁴⁰

In Zayers Modell wird unter dem Begriff Self-Justification eine Veränderung der Ziele des Entscheiders verstanden.¹⁴¹ Wurden zuvor Unternehmensziele verfolgt, steht nun der Schutz des Selbstbilds an erster Stelle. Ist der Verantwortliche bereit, seinen Entscheidungsfreiraum für den Schutz seines Selbstbilds auszunutzen, so führt die Veränderung der Zielsetzung zur Fortführung des Projekts, obwohl der Abbruch ökonomisch vorteilhafter wäre. Damit ist der Self-Justification-Effekt beschrieben. Der Effekt als solcher ist jedoch für eine empirische Erhebung kaum valide in einer einzelnen Variablen zu operationalisieren. Die Stärke der verwendeten Kovarianzanalyse besteht nämlich gerade darin, den Effekt (d. h. die Wirkung) *nicht* direkt zu erfragen, sondern zwei Variablen zu erfassen, aus deren Kovarianz der Effekt geschätzt werden kann. Um die Stringenz der kovarianzbasierten Messungen beizubehalten, wird der Effekt in der vorliegenden Arbeit indirekt gemessen. Zu diesem Zweck werden die Respondenten nach der Ausprägungsstärke der Faktoren gefragt, die den Effekt hervorrufen. Diese Faktoren bestehen in der Wahrnehmung von Druck, das Gesicht zu wahren und das Image schützen zu müssen, wie oben erläutert wurde. Um die Messung der Variablen möglichst exakt widerzugeben, wird das empirisch erhobene Konstrukt im Folgenden als „Self-Justification-Druck“ bezeichnet.

H₃: Der Self-Justification-Druck wirkt direkt positiv auf die Projekteskalation.

Zayer identifiziert drei Ansatzpunkte zur Vermeidung des Self-Justification-Effekts. Alle drei Maßnahmen beziehen sich in ihrer Wirkung gleichermaßen auf die Reduzierung des Self-

¹³⁵ Vgl. Staw 1976.

¹³⁶ Vgl. Abbildung 2.1 auf Seite 13.

¹³⁷ Vgl. Kiesler/Sakumura 1966.

¹³⁸ Vgl. Festinger 1957.

¹³⁹ Für Analysen der empirischen Daten unter expliziter Berücksichtigung der Theorie der kognitiven Dissonanz vgl. Abschnitt 4.4.1.3 auf Seite 152.

¹⁴⁰ Vgl. Brockner/Rubin 1985, Brockner/Rubin/Lang 1981, S. 71 f., Brockner et alii 1982, S. 263, Simonson/Staw 1992, S. 419, Staw 1976, S. 42 und Staw 1981.

¹⁴¹ Vgl. Zayer 2007, S. 115 ff.

Justification-Drucks. Die Ansatzpunkte sind der Grad der wahrgenommenen persönlichen Verantwortung, die Stärke der inneren Bedrohung des Selbstbilds und die Stärke der Bedrohung des Selbstbilds durch Dritte.¹⁴²

Reduktion der wahrgenommenen persönlichen Verantwortung

Während frühere Literatur vorschlägt, für den Erfolg oder Misserfolg eines Projekts *einen* Akteur verantwortlich zu machen,¹⁴³ kommt die neuere Eskalationsforschung zu dem Schluss, dass eine Aufteilung der Ergebnisverantwortung die Eskalationstendenzen verringern kann. Um eine Reduktion der wahrgenommenen persönlichen Verantwortung des Projektverantwortlichen zu erreichen, können andere Personen (beispielsweise der Controller) einen Teil der Verantwortung für das Scheitern des Projekts mitübernehmen.¹⁴⁴

H_{M9}: Die Mitverantwortung anderer Personen wirkt direkt negativ auf den Self-Justification-Druck.

Verringerung der internen Bedrohung des Selbstbilds

Simonson/Staw weisen in einem Experiment nach, dass Entscheidungsverantwortliche signifikant weniger Ressourcen in ein scheiterndes Projekt investieren, wenn sie nach der Qualität ihres Entscheidungsprozesses bewertet werden und nicht nur nach den Folgen bzw. dem Ergebnis der Entscheidung.¹⁴⁵ Dementsprechend erscheint die Verwendung der *Prozessqualität* als zusätzliches Kontrollobjekt sinnvoll, um die Eskalationstendenz zu reduzieren.¹⁴⁶

H_{M10}: Das Bewerten der Qualität der Entscheidungsprozesse wirkt direkt negativ auf den Self-Justification-Druck.

¹⁴² Vgl. Zayer 2007, S. 117 ff. und 185 ff.

¹⁴³ Vgl. Jones/McLean 1970, S. 8.

¹⁴⁴ Vgl. Heng/Tan/Wei 2003. Ähnlich Staw/Ross 1987b, S. 73.

¹⁴⁵ Vgl. Simonson/Staw 1992, S. 424.

¹⁴⁶ Vgl. hierzu auch Keil/Robey 1999 und Kirby/Davis 1998. Eine Bewertung nur anhand des Ergebnisses ist schon allein deshalb kritisch zu betrachten, da die beurteilenden Personen im Nachhinein ihre eigene Fähigkeit überschätzen, das Resultat einer Handlung vorherzusehen; vgl. beispielsweise Brown/Solomon 1993.

Verringerung der Bedrohung des Selbstbilds durch Dritte

Simonson/Staw führen experimentell den Nachweis, dass eine Reduzierung des Bedrohungsgefühls zu einer geringeren Eskalation führt.¹⁴⁷ Als Maßnahme zur Verringerung der Bedrohung bietet sich beispielsweise an, dass der Controller explizit kommuniziert, dass weder Ansehen noch Karriere des Entscheiders durch einen Abbruch gefährdet sind.¹⁴⁸ Keil/Robey bezeichnen dies als „Organizational tolerance for failure“.¹⁴⁹

H_{M11}: Der Schutz des Ansehens wirkt direkt negativ auf den Self-Justification-Druck.

2.3.4 Bewertungsfehler

Konnte eine Veränderung des Zielsystems vermieden werden, liegt nach dem Modell von Zayer die letzte Hürde beim Entscheidungsprozess im Auftreten von Bewertungsfehlern.¹⁵⁰ Insbesondere zwei Bewertungsfehler können die Eskalationstendenz verstärken: der Sunk-Cost-Effekt und (übermäßiger) Optimismus.

2.3.4.1 Sunk-Cost-Effekt

Als Sunk Costs werden Investitionen (Geld, Zeit, Anstrengung etc.) bezeichnet, die in der Vergangenheit getätigt wurden und nicht wieder zurückgewonnen werden können.¹⁵¹ Menschen wollen in der Regel Ressourcen nicht unnötig verschwenden.¹⁵² Der Abbruch eines scheiternden Projekts kann jedoch den Eindruck von Verschwendung erwecken, wenn mit dem Abbruch bereits getätigte Investitionen unwiederbringlich verlorengehen. Um diesen Verlust zu vermeiden, neigen Menschen dazu, in ein scheiterndes Projekt weitere Ressourcen zu investieren, was als *Sunk-Cost-Effekt* bezeichnet wird.¹⁵³ Dieses Verhalten steht im Widerspruch zur

¹⁴⁷ Vgl. Simonson/Staw 1992, S. 424.

¹⁴⁸ Vgl. Heng/Tan/Wei 2003.

¹⁴⁹ Keil/Robey 1999, S. 68.

¹⁵⁰ Vgl. Zayer 2007, S. 135 ff.

¹⁵¹ Vgl. Hammond/Keeney/Raiffa 2006, S. 122. Sunk Costs spielen in vielen wirtschaftlichen Zusammenhängen eine Rolle. Für eine Einordnung in den Resourced-Based-View und den Zusammenhang zwischen Sunk Costs und Marktversagen vgl. beispielsweise Mahoney/Pandian 1992, S. 369 f. Im Rahmen dieser Untersuchung wird jedoch auf die *psychologische* Wirkung der Sunk Costs im Projektkontext fokussiert.

¹⁵² Vgl. Conlon/Leatherwood 1989, S. 48.

¹⁵³ Vgl. Dilts/Pence 2006b, S. 379.

präskriptiven Entscheidungstheorie, nach der nur die zukünftigen Zahlungsströme in die Bewertung einer Allokationsentscheidung einfließen dürfen.¹⁵⁴ Der Sunk-Cost-Effekt, der treffend mit der Wendung „throwing good money after bad“ charakterisiert wird,¹⁵⁵ führt zu einer Verstärkung der Projekteskalation.

H₄: Der Sunk-Cost-Effekt wirkt direkt positiv auf die Projekteskalation.

In der experimentellen Forschung wurde festgestellt, dass sowohl die absolute Höhe¹⁵⁶ der Sunk Costs als auch ihre relative Höhe¹⁵⁷ die Eskalationstendenz verstärken können. Basierend auf der Literatur schlägt Zayer eine Reihe von Maßnahmen vor, die sowohl die Wirkung der absoluten als auch der relativen Sunk Costs auf den Entscheidungsprozess verhindern sollen.¹⁵⁸ Diese werden im Folgenden näher erläutert.

Vermeidung der irrationalen Bewertung von relativen und absoluten Sunk Costs

Um Bewertungsfehler durch die Sunk Costs zu verhindern, wird in der Literatur vorgeschlagen, die Sunk Costs explizit auszuweisen und darauf hinzuweisen, dass diese nicht einbezogen werden dürfen.¹⁵⁹ Eine Verbesserung der Bewertung ist des Weiteren durch den Einsatz rationaler Bewertungsinstrumente, beispielsweise der Kapitalwertmethode, zu erwarten.¹⁶⁰ Darüber hinaus kann der Verantwortliche Unterstützung von Methodenexperten (beispielsweise Controllern) erhalten.¹⁶¹ In einem Laborexperiment konnte gezeigt werden, dass die Möglichkeit,

¹⁵⁴ Vgl. exemplarisch Hummel/Männel 1986, S. 117. Während die *deskriptive* Entscheidungstheorie beschreibt, wie in der Realität Entscheidungen getroffen werden, zeigt die *präskriptive oder normative* Entscheidungstheorie, wie Entscheidungsprobleme rational gelöst werden können; vgl. beispielsweise Laux 2005, S. 2, oder Jungermann/Pfister/Fischer 1998, S. XI. Die Entscheidungstheorie will dem „Entscheider nicht dogmatisch vorschreiben, was er tun soll“, sondern „ihm helfen, seine eigenen Zielvorstellungen in ein widerspruchsfreies ‚Zielsystem‘ zu überführen und dann eine Entscheidung zu treffen, die mit diesem Zielsystem im Einklang steht. Die Entscheidungstheorie nimmt – im Gegensatz zur Ethik – keine Wertung der Zielvorstellungen des Entscheiders vor; sie nimmt sie als gegeben an, ohne sie beeinflussen zu wollen.“; Laux 2005, S. 3. „Die Betriebswirtschaftslehre wird [. . .] häufig als spezielle (oder angewandte) Entscheidungstheorie bezeichnet.“; Laux 2005, S. 2.

¹⁵⁵ Vgl. Garland/Sandefur/Rogers 1990, Ghosh 1995 und Guler 2007.

¹⁵⁶ Vgl. Arkes/Blumer 1985, S. 123 ff., und Heath 1995, S. 47 ff.

¹⁵⁷ Vgl. Garland/Sandefur/Rogers 1990, Keil et alii 1995, Moon 2001a und Moon 2001b.

¹⁵⁸ Vgl. Zayer 2007, S. 200 ff.

¹⁵⁹ Vgl. Conlon/Garland 1993, S. 412, und Conlon/Leatherwood 1989, S. 51 ff.; diese Vorschläge wurden bislang noch nicht in Experimenten geprüft.

¹⁶⁰ Vgl. Ghosh 1997, Newman/Sabherwal 1996 und Northcraft/Wolf 1984.

¹⁶¹ Vgl. Weber 2002, S. 32.

in andere Projekte zu investieren, signifikant die Allokation von Ressourcen zu einem negativ verlaufenden, eskalierenden Projekt reduziert.¹⁶² Daraus ergibt sich die Maßnahme, Opportunitätskosten-Informationen auszuweisen, um das zukunftsgerichtete Denken zu fördern.¹⁶³

H_{M12}: Der explizite Ausweis der Sunk Costs wirkt direkt negativ auf den Sunk-Cost-Effekt.

H_{M13}: Die Kennzeichnung der Irrelevanz von Sunk Costs wirkt direkt negativ auf den Sunk-Cost-Effekt.

H_{M14}: Angemessene Verfahren zur Projektbewertung wirken direkt negativ auf den Sunk-Cost-Effekt.

H_{M15}: Expertenunterstützung wirkt direkt negativ auf den Sunk-Cost-Effekt.

H_{M16}: Der Hinweis auf Alternativen wirkt direkt negativ auf den Sunk-Cost-Effekt.

2.3.4.2 Optimismus

Eine rationale Bewertung des Projekts kann auch durch (übermäßigen) Optimismus beeinträchtigt werden.¹⁶⁴ Optimismus kann beispielsweise dazu führen, dass Erfolgsaussichten systematisch überschätzt und Bedrohungen systematisch unterschätzt werden.¹⁶⁵ Das Auftreten von Optimismus ist durch zahlreiche Studien belegt worden.¹⁶⁶ Im Bezug auf Projektentscheidungen ist die logische Konsequenz einer durch Optimismus verzerrten Bewertung der Erfolgswahrscheinlichkeiten das Festhalten an scheiternden Projekten.

H₅: Optimismus wirkt direkt positiv auf die Projekteskalation.

Das Ausmaß der optimistischen Verzerrung wird dabei vor allem durch drei Faktoren determiniert: die Kontrollillusion, die früheren Erfahrungen des Verantwortlichen und Overconfidence.¹⁶⁷ Overconfidence wird, als moderierender Effekt, in Abschnitt 2.3.5.3 auf Seite 43 beschrieben. Der Einfluss der Kontrollillusion und der früheren Erfahrungen wird im Folgenden näher erläutert.

¹⁶² Vgl. Keil/Truex III/Mixon 1995, S. 375 ff., Keil et alii 1995, S. 74 ff., und McCain 1986, S. 282.

¹⁶³ Vgl. auch Keil/Robey 1999 und Northcraft/Neale 1986.

¹⁶⁴ Vgl. Zayer 2007, S. 142.

¹⁶⁵ Vgl. beispielsweise Bird 2005, Weinstein 1980, Weinstein 1987 und Weinstein 1996.

¹⁶⁶ Vgl. beispielsweise Harris/Middleton 1994, March/Shapira 1987, Svenson 1981 und Weinstein 1987.

¹⁶⁷ Vgl. Zayer 2007, S. 143.

Reduzierung der Kontrollillusion

Der Begriff Kontrollillusion beschreibt das Phänomen, dass Individuen häufig überschätzen, wie stark ihre Handlungen ein Ergebnis beeinflussen können. Dies führt dazu, dass sie zu große Risiken eingehen.¹⁶⁸ Um diesen Effekt zu verhindern, wird in der Literatur vorgeschlagen, Schätzungen der zukünftigen Kosten und Erlöse des Projekts durch Analogien zu anderen Projekten vorzunehmen.¹⁶⁹

H_{M17}: Analogien mit anderen Projekten wirken direkt negativ auf den Optimismus.

Korrektur der durch die individuellen Erfahrungen geprägten optimistischen Verzerrung

Eine seit langem bekannte Technik zur Reduzierung optimistischer Verzerrungen ist die Methode des *Advocatus Diaboli*.¹⁷⁰ Darunter wird eine Reihe von Techniken verstanden, die einen strukturierten Konflikt in den Prozess der Entscheidungsfindung einführen.¹⁷¹ Schwenk weist in einem Experiment nach, dass diese Maßnahme zu einer Reduzierung der Eskalation des Commitments beiträgt.¹⁷²

H_{M18}: Der Einsatz eines *Advocatus Diaboli* wirkt direkt negativ auf den Optimismus.

2.3.5 Moderierende Faktoren

Das Modell von Zayer beinhaltet drei moderierende Faktoren: den Zeitpunkt der Wahrnehmung, die Bereitschaft, Freiraum auszunutzen, und die Overconfidence.

2.3.5.1 Zeitpunkt der Wahrnehmung

Als moderierender Faktor zwischen Wahrnehmungsschwelle und Wahrnehmungsdefizit wird der Zeitpunkt der Wahrnehmung betrachtet.¹⁷³ Dieser Effekt ist wie folgt zu beschreiben: Wer-

¹⁶⁸ Vgl. Boulding/Morgan/Staelin 1997, S. 175, Camerer 1999, Du et alii 2007, S. 279, und Strickland/Lewicki/Katz 1966.

¹⁶⁹ Vgl. Lovallo/Kahneman 2003.

¹⁷⁰ Vgl. beispielsweise Cosier 1981 und Schweiger/Sandberg/Ragan 1986.

¹⁷¹ Vgl. Schwenk 1988, S. 769.

¹⁷² Vgl. Schwenk 1988, S. 777.

¹⁷³ Vgl. Abbildung 2.2 auf Seite 22 und Zayer 2007, S. 110 f.

den die negativen Informationen früh auffällig,¹⁷⁴ reduziert dies die Projekteskalation. Werden die Probleme jedoch erst sehr spät augenfällig, so können sie eskalierend wirken.¹⁷⁵ Anders ausgedrückt, wenn die negativen Informationen erst sehr spät auftreten, so verstärken sie eine Eskalation, so dass in diesem Fall eine höhere Wahrnehmungsschwelle die Eskalation abschwächen kann.

H₆: Der Zeitpunkt der Wahrnehmung moderiert die Wirkung der Wahrnehmungsschwelle auf die Projekteskalation negativ.

Für den moderierenden Faktor Zeitpunkt der Wahrnehmung werden von Zayer keine separaten Maßnahmen zur Reduzierung aufgeführt.

2.3.5.2 Bereitschaft, Freiraum auszunutzen

Gemäß Zayers Modell wird die Wirkung des Self-Justification-Effekts auf die Eskalationstendenz durch die Bereitschaft des Verantwortlichen, den Entscheidungsfreiraum auszunutzen, moderiert.¹⁷⁶ Der moderierende Faktor wird bei Zayer als „Bereitschaft, Freiraum auszunutzen“ bezeichnet. Die Maßnahmen, die zur Vermeidung dieses Faktors angeführt werden, beziehen sich jedoch nicht nur auf die Bereitschaft, sondern zum Teil auch auf die „Verringerung der Entscheidungsfreiheit des Projektleiters“.¹⁷⁷ Es erscheint deshalb präziser, das Konstrukt im Folgenden als „nicht hinterfragter Freiraum“ zu bezeichnen.

Ein hoher, nicht hinterfragter Freiraum des Verantwortlichen begünstigt, dass der Entscheidungsverantwortliche dem Self-Justification-Druck nachgibt und das Projekt fortführt, selbst wenn dies den ökonomischen Interessen des Unternehmens entgegensteht.¹⁷⁸

H₇: Der nicht hinterfragte Entscheidungsfreiraum des Entscheidungsverantwortlichen moderiert die Wirkung von Self-Justification-Druck auf die Projekteskalation positiv.

¹⁷⁴ Der Fall, dass auch positive Informationen durch die Wahrnehmungsschwelle unbemerkt bleiben können, ist für das Verständnis verspäteter Projektabbrüche und damit für die vorliegende Arbeit nicht relevant. Studien, die sich verfrühten Projektabbrüchen widmen, sollten hingegen diese Möglichkeit berücksichtigen.

¹⁷⁵ Vgl. Brockner et alii 1982. Dieser Effekt könnte durch den Self-Justification- oder den Sunk-Cost-Effekt erklärt werden; vgl. Zayer 2007, S. 110.

¹⁷⁶ Vgl. Abbildung 2.2 auf Seite 22 und Zayer 2007, S. 133 f.

¹⁷⁷ Zayer 2007, S. 195.

¹⁷⁸ Vgl. Zayer 2007, S. 195.

Zayer führt mehrere Maßnahmen auf, mit denen die Eskalation auch bei bestehender Bedrohung des Selbstbilds verhindert werden kann.¹⁷⁹ Eine Möglichkeit besteht in der Offenlegung der Ziele des Entscheidungsverantwortlichen und einer *Diskussion* der Erreichbarkeit der Ziele bei Projektdurchführung.¹⁸⁰ Ebenfalls experimentell nachgewiesen ist die eskalationssenkende Wirkung der Festlegung einer *öffentlich bekannten Abbruchregel* zu Beginn des Projekts.¹⁸¹ Keil/Robey bezeichnet diese Maßnahme als „the presence of publicly stated limits“.¹⁸²

H_{M19}: Die Bekanntgabe der Ziele, die mit dem Projekt verfolgt werden, wirkt direkt negativ auf den nicht hinterfragten Freiraum.

H_{M20}: Die Diskussion der Ziele, die mit dem Projekt verfolgt werden, wirkt direkt negativ auf den nicht hinterfragten Freiraum.

H_{M21}: Die frühe Festlegung von Abbruchkriterien wirkt direkt negativ auf den nicht hinterfragten Freiraum.

2.3.5.3 Overconfidence

Das Metawissen um die Zuverlässigkeit des eigenen Wissens ist für rationale Entscheidungen unabdingbar.¹⁸³ Es beinhaltet das Wissen um die Unsicherheit unserer Schätzungen und Vorhersagen.¹⁸⁴ Unter Overconfidence wird diesbezüglich bei Zayer und auch in der vorliegenden Arbeit das Überschätzen der Güte der eigenen Prognosen für den Projekterfolg verstanden.¹⁸⁵ Individuen neigen dazu, ihre eigenen Schätzungen für korrekter zu halten, als sie tatsächlich sind.¹⁸⁶ Eine empirische Untersuchung des Phänomens im Unternehmenskontext von Russo/Schoemaker kam zu dem Ergebnis, dass weniger als ein Prozent der über 2000 Befragten *nicht* overconfident war.¹⁸⁷

¹⁷⁹ Vgl. Zayer 2007, S. 195 ff.

¹⁸⁰ Vgl. Teger 1980.

¹⁸¹ Vgl. Boulding/Morgan/Staelin 1997, Brockner/Shaw/Rubin 1979, Keil/Robey 1999, Simonson/Staw 1992, S. 424, und Teger 1980.

¹⁸² Keil/Robey 1999, S. 68.

¹⁸³ Vgl. Russo/Schoemaker 1992, S. 7.

¹⁸⁴ Vgl. Russo/Schoemaker 1992, S. 8.

¹⁸⁵ Vgl. Zayer 2007, S. 148.

¹⁸⁶ Vgl. Lichtenstein/Fischhoff 1977, Lichtenstein/Fischhoff/Phillips 1982, Oskamp 1965 und Tversky/Kahneman 1974.

¹⁸⁷ Vgl. Russo/Schoemaker 1992, S. 9.

Menschen neigen dazu, Zuversicht (confidence) mit Kompetenz gleichzusetzen. Dementsprechend kann davon ausgegangen werden, dass viele Entscheider Hemmungen haben, beispielsweise in einem wichtigen Meeting klarzustellen, dass sie sich *nicht* sicher sind.¹⁸⁸ Der Hang zur Overconfidence nimmt dabei mit zunehmender Erfahrung ab, bleibt aber auch bei Experten noch deutlich bestehen. Russo/Schoemaker führen dies darauf zurück, dass aus der Erfahrung nicht ausreichend gelernt wird.¹⁸⁹

Im Kontext risikoreicher Projekte verstärkt Overconfidence Probleme wie das Unterschätzen von Risiko sowie das Überschreiten von Deadlines und Budgets.¹⁹⁰ Es ist deshalb davon auszugehen, dass sich der Optimismus eines Entscheidungsträgers um so stärker auf die Projekteskalation auswirkt, je stärker seine Overconfidence ausgeprägt ist.¹⁹¹

H₈: Overconfidence moderiert die Wirkung von Optimismus auf die Projekteskalation positiv.

Für den moderierenden Faktor Overconfidence führt Zayer keine separaten Maßnahmen zur Reduzierung auf.

¹⁸⁸ Vgl. Russo/Schoemaker 1992, S. 16.

¹⁸⁹ Vgl. Russo/Schoemaker 1992, S. 7 ff.

¹⁹⁰ Vgl. Russo/Schoemaker 1992, S. 9.

¹⁹¹ Eine der wenigen Berufsgruppen, bei der Overconfidence *nicht* signifikant messbar ist, sind Wirtschaftsprüfer; vgl. Russo/Schoemaker 1992, S. 10. Russo/Schoemaker schreiben dies dem Umstand zu, dass der Berufsethos von Wirtschaftsprüfern konservativen Einschätzungen starken Wert beimisst und Wirtschaftsprüfer gelernt haben, Overconfidence zu kompensieren. In der Tat konnte bei Wirtschaftsprüfern sogar eine Underconfidence nachgewiesen werden; vgl. Tomassini et alii 1982.

3 Methodik und Durchführung der empirischen Studie

3.1 Diskussion der Datenerhebung und Datenbasis

3.1.1 Auswahl der Methode der Datenerhebung

Nachdem die Forschungsfragen, die zu prüfenden Hypothesen und das theoretische Modell erarbeitet wurden, soll nun die Wahl der Datenerhebungsmethode begründet werden. Die Forschungsfragen der vorliegenden Arbeit¹ sind vor allem durch zwei Merkmale geprägt, die relevant für die Untersuchungsmethodik sind. Das ist zum einen (1) die Fokussierung auf den realen Unternehmenskontext und zum anderen (2) die Frage nach der relativen Einflussstärke der Faktoren.

(1) Mit dem Fokus auf den realen Unternehmenskontext scheiden sach- und formalanalytische Verfahren sowie Simulationen aus. Auch Experimente² sind wenig geeignet. Insbesondere stellt

¹ Vgl. Abschnitt 1.3 auf Seite 7.

² Es erscheint angebracht, die Verwendung von Laborexperimenten etwas ausführlicher zu diskutieren, da sie in der Behavioral-Accounting-Forschung die häufigste Datenerhebungsmethode darstellen. Einen exemplarischen Einblick in die vorherrschende Methodik in der Behavioral-Accounting-Forschung bietet beispielsweise der Artikel von Meyer/Rigsby 2001. Darin untersuchen Meyer/Rigsby 134 Beiträge, die zwischen 1989 und 1998 in der Zeitschrift Behavioral Research in Accounting (BRIA) veröffentlicht wurden. Von diesen basieren 68 auf experimentellen Methoden; vgl. Meyer/Rigsby 2001, S. 259 ff. Diese Beobachtungen treffen auch speziell für die Eskalationsforschung zu. Ein großer Teil der Eskalationsstudien verwendet dabei Studenten als Subjekte der Experimente; vgl. Simonson/Staw 1992, S. 421.

sich die Frage nach der externen Validität von Laborexperimenten.³ Zwar können Experimente Personen aus der Zielgruppe untersuchen, allerdings erzwingt die notwendige Kontrolle eine starke Vereinfachung des Entscheidungsproblems. Deshalb gilt als ein Hauptproblem bei Experimenten die Übertragbarkeit der Erkenntnisse auf reale Situationen. Als Resultat seiner Modellierung von Entscheidungssituationen als komplexe Spiele zieht beispielsweise Dreze die Schlussfolgerung: "It seems highly dubious that observations about such trivial decisions could by extrapolation throw any light on the behavior of similar persons when faced with truly important problems."⁴ Für eine Untersuchung im Unternehmenskontext eignen sich deshalb andere Methoden besser, beispielsweise Fallstudien auf der Basis von Interviews oder Feldstudien sowie großzahlige Umfragen.

(2) Da die Forschungsfragen sich auf die relative Stärke der Faktoren beziehen, sind Fallstudien wenig geeignet, da sie lediglich *qualitative* Aussagen über die Stärke der Faktoren geben können. Bei Tiefeninterviews und der Erstellung von Fallstudien ergibt sich für die vorliegende Fragestellung zudem insbesondere das Problem, dass die Befragten gehemmt sein könnten, von negativen Projektverläufen zu berichten. Darüber hinaus bliebe die Übertragbarkeit der Ergebnisse auf die Gesamtheit ungeklärt. Nur eine großzahlige Umfrage bietet die Möglichkeit, mit verlässlicher Genauigkeit auf die *Häufigkeit* bestimmter Phänomene in der relevanten Gesamtpopulation zu schließen.⁵ Gleichzeitig ermöglicht die großzahlige Umfrage, die Verläufe der Wirkungen zwischen vielen Variablen zu untersuchen.⁶

Eine großzahlige Befragung stellt demnach für die genannten Forschungsfragen die am besten geeignete Methode der Datenerhebung dar. Die Fragen können dabei entweder durch einen Interviewer (beispielsweise bei Telefoninterviews) gestellt oder selbstständig bearbeitet werden (beispielsweise in Form von Fragebögen). Zahlreiche Forscher haben sich mit der Frage beschäftigt, inwieweit Charakteristika des *Interviewenden* die Antworten des Befragten beeinflussen.⁷ Diese Beeinflussung der Ergebnisse durch den Interviewenden bzw. den Experimentator wird als *Experimenter Effect* bezeichnet und spielt gerade in der verhaltenswissenschaftlichen Forschung eine wichtige Rolle.⁸ Durch eine internetbasierte oder postalische Befragung

³ Vgl. hierzu ausführlich Abschnitt 1.2.1 auf Seite 2.

⁴ Dreze 1958, S. 48; zitiert nach Stedry 1960, S. 55.

⁵ Vgl. Dillman 2000, S. 9.

⁶ Vgl. Maines/Salamon/Sprinkle 2006, S. 97.

⁷ Vgl. beispielsweise Campbell/Holder/France 2006, Kurzban/DeScioli/O'Brien 2007 und Rosenthal 1976, S. 112.

⁸ Vgl. Rosenthal 1976.

kann eine Vielzahl von Interviewer-Biases ausgeschlossen werden,⁹ die durch persönlichen Kontakt hervorgerufen werden können.¹⁰ Umfangreiche empirische Befunde bestätigen beispielsweise, dass Menschen bei selbstständig auszufüllenden Fragebögen ehrlicher antworten als in persönlichen oder telefonischen Interviews.¹¹ Dies ist insbesondere bei Fragen zu sozial unerwünschtem Verhalten, die auch in dieser Befragung verwendet werden müssen, ein wichtiger Faktor. Entsprechend werden für die empirische Erhebung in dieser Studie selbstständig auszufüllende Fragebögen verwendet.

Abschließend gilt noch zu klären, ob der Fragebogen den Respondenten postalisch oder online zur Verfügung gestellt werden sollte. Einzelne Studien konstatierten als Nachteil computergestützter Befragungen im Vergleich zu Befragungen in Papierform, dass der Prozentsatz zufälliger Daten höher ist.¹² Diese entstehen durch Personen, die die Fragen nicht gelesen, sondern zufällige Antworten gegeben haben. Um diese Fehlerquelle zu kontrollieren, wurden verschiedene Maßnahmen durchgeführt, die in Abschnitt 4.1.1.2 auf Seite 100 näher erläutert werden.

In Bezug auf die inhaltlichen Aussagen der Respondenten wurden keine wesentlichen Unterschiede zwischen Fragebögen im Internet und Papierfragebögen gefunden.¹³ Die Vorteile einer Online-Befragung bestehen beispielsweise in der Möglichkeit, Filtervariablen einzusetzen, was auch in der vorliegenden Untersuchung genutzt wurde. Außerdem entfällt die potentiell fehleranfällige Übertragung der Papierfragebögen in eine Datenbank.¹⁴ Darüber hinaus lässt sich bei einer Online-Befragung die Bearbeitungszeit messen, was bei postalisch übermittelten Fragebögen nicht möglich ist. Die Messung der Bearbeitungszeit kann bei der Identifikation flüchtig ausgefüllter Fragebögen von Nutzen sein.¹⁵

⁹ Für einen tabellarischen Überblick über Maßnahmen zur Reduzierung und Kontrolle des Experimentier- bzw. Interviewer-Bias vgl. Rosenthal 1976, S. 402; für Erläuterungen der Maßnahmen vgl. Rosenthal 1976, S. 331 ff.

¹⁰ Für einige interessante Beispiele nicht intendierter Beeinflussung von Befragten vgl. beispielsweise Rosenthal 1976, S. 112 ff.

¹¹ Vgl. beispielsweise Aquilino 1994, Dillman 1991 und Fowler/Roman/Di 1998.

¹² Vgl. Beach 1989, S. 102. Die computergestützten Fragebögen im Experiment von Beach 1989 wurden den Teilnehmern nicht online präsentiert. Da Online-Befragungen jedoch notwendigerweise über den Computer erfolgen, erscheint es angebracht, auch die Erkenntnisse computergestützter Experimente zu berücksichtigen, die offline durchgeführt wurden.

¹³ Vgl. beispielsweise Griffis/Goldsby/Cooper 2003, S. 246.

¹⁴ Vgl. Griffis/Goldsby/Cooper 2003, S. 250.

¹⁵ Vgl. Abschnitt 4.1.1.2 auf Seite 100.

In Bezug auf die Antwortrate kommt eine Untersuchung von Boyer et alii zu dem Schluss, dass keine signifikanten Unterschiede zwischen Papierfragebögen und Internetfragebögen auftreten.¹⁶ Eine Studie von Griffis/Goldsby/Cooper hingegen ergab eine signifikant *höhere* Rücklaufquote für eine Online-Befragung (14,3 Prozent) im Vergleich zu einer papierbasierten Befragung (10,1 Prozent).¹⁷ Gibt man den Respondenten die Möglichkeit, zwischen Papierfragebogen und Internetfragebogen zu wählen, fällt die Wahl signifikant häufiger auf die Internetvariante. Einen Fragebogen online auszufüllen, scheint somit für die Mehrzahl der potentiellen Respondenten die attraktivere Alternative zu sein.¹⁸ Die Verwendung eines webbasierten Fragebogens erscheint demzufolge als die beste Lösung.¹⁹ Die Online-Befragung wird auch bei vergleichbaren Studien bevorzugt²⁰ und soll deshalb für die vorliegende Untersuchung ebenfalls eingesetzt werden.

3.1.2 Auswahl der Zielpopulation und der Stichprobe

Nach der Auswahl der Methode der Datenerhebung ist zu entscheiden, welche Subjekte befragt werden sollen. Empirische Studien haben gezeigt, dass die Funktion des Respondenten einen höheren Einfluss auf die Varianz der Ergebnisse ausüben kann als der eigentliche Untersuchungsgegenstand (trait).²¹ Auch die Position in der Hierarchie wirkt sich stark auf die Ergebnisse aus.²² Dies lässt Studien, die lediglich auf einem Informanten aufbauen, fragwürdig erscheinen.²³ Um diesem Problem entgegenzuwirken, wurden zwei verschiedene Gruppen von Funktionsträgern ausgewählt, die ein hohes Potential haben, sich mit abgebrochenen Projekten

¹⁶ Vgl. Boyer et alii 2002.

¹⁷ Vgl. Griffis/Goldsby/Cooper 2003, S. 243.

¹⁸ Vgl. Griffis/Goldsby/Cooper 2003, S. 248.

¹⁹ Für einen ausführlicheren Vergleich im Hinblick auf eine Vielzahl von Kriterien vgl. Griffis/Goldsby/Cooper 2003, S. 252.

²⁰ Vgl. Dilts/Pence 2006b, S. 385.

²¹ Vgl. Ernst 2003a, S. 14 f.

²² Vgl. Ernst 2003a, S. 15 f.

²³ Vgl. hierzu den Abschnitt 3.3.3 auf Seite 65, der sich ausführlicher mit dem Informant-Bias auseinandersetzt.

auszukennen. Zum einen wurden Controller gewählt,²⁴ zum anderen Führungskräfte in technischen Bereichen.²⁵

Für die Beschaffung der Adressen wurde die Hoppenstedt-Firmendatenbank verwendet.²⁶ Daraus wurden Personen selektiert, die den folgenden Kriterien entsprachen:

Funktion: Controlling oder technische Unternehmensführung oder Leitung Forschung und Entwicklung.

Position: In Bezug auf die Position innerhalb der Unternehmenshierarchie wurden alle Positionen zugelassen, mit einer Ausnahme: Für Großunternehmen wurden Personen auf der obersten Hierarchieebene ausgeschlossen.²⁷ Der Grund für diese Einschränkung ist die Überlegung, dass Personen in diesen Positionen bei Großunternehmen mit deutlich geringerer Wahrscheinlichkeit den Fragebogen ausfüllen als Personen, die niedriger in der Unternehmenshierarchie stehen.²⁸

Firmentyp: Es wurden alle Firmentypen selektiert, die als Firma, Konzern oder Bankzentrale klassifiziert waren. Damit wurden Behörden, Verbände und Bankfilialen ausgeschlossen. Dieser Ausschluss basiert auf der Überlegung, dass diese Firmentypen eine deutlich andere Ausprägung des Controllings aufweisen und/oder sich die Art von Projekten und deren Zielsetzungen deutlich von anderen (privatwirtschaftlichen) Unternehmen unterscheiden können.

Branche: In Bezug auf die Branche wurden keine Einschränkungen festgelegt, da das theoretische Modell als branchenübergreifend gültig befunden wurde.²⁹ Auf Grund des bran-

²⁴ Keil/Robey heben die Eignung von Auditoren hervor, da bei ihnen von einer höheren Objektivität ausgegangen werden kann als bei den direkt am Projekt beteiligten Personen; vgl. Keil/Robey 1999, S. 71. Analog kann argumentiert werden, dass Controller als kritischer Counterpart eine höhere Objektivität aufweisen können als Manager.

²⁵ Die statistische Überprüfung des Einflusses der Funktion der Befragten wird in Abschnitt 4.6.3 auf Seite 198 dargelegt.

²⁶ Es wurde die Version 2.1/2007 (Relaunch Update), Stand: Juni 2007, verwendet; vgl. Hoppenstedt Firmendatenbank 2007.

²⁷ Zur obersten Hierarchieebene wurden folgende Positionsbezeichnungen gerechnet: geschäftsführendes Vorstandsmitglied, Geschäftsführung, Geschäftsleiter, Geschäftsleitung, Hauptgeschäftsführer, Verwaltungsrat, Vorsitzender, Vorstand und Vorstandsvorsitzender.

²⁸ Die Auswirkung von Respondentenmerkmalen auf die Rücklaufquote stellt eine eigene Strömung in der Nonresponse-Forschung dar; vgl. Bosnjak/Tuten/Wittmann 2005, S. 490.

²⁹ Vgl. hierzu die Diskussion in Abschnitt 2.2.2.1 auf Seite 23.

chenübergreifenden Ansatzes ist davon auszugehen, dass die Ergebnisse hohe Allgemeingültigkeit, d. h. eine hohe externe Validität, aufweisen.³⁰

Unternehmensgröße: Da bei kleinen Unternehmen im Regelfall nicht vom Einsatz eines ausgefeilten Projektcontrollings auszugehen ist, wurden nur mittlere und große Unternehmen gemäß der Definition der Kommission der Europäischen Gemeinschaften ausgewählt, d. h. Unternehmen mit 50 Beschäftigten oder mehr *oder* einem Jahresumsatz bzw. einer Jahresbilanzsumme von mehr als zehn Mio. Euro.³¹

Beim Ziehen einer Stichprobe besteht die Gefahr des Sampling Errors.³² Darunter wird die Abweichung der Daten einer Stichprobe von den Daten der relevanten Gesamtpopulation verstanden.³³ Der Effekt des Sampling Errors ist jedoch für die vorliegende Studie als eher gering einzuschätzen. So beträgt beispielsweise der Sampling Error nur \pm zehn Prozent, wenn aus den Antworten von 100 Befragten auf die gesamten Einwohner der USA geschlossen wird.³⁴ Da im vorliegenden Fall ein deutlich höherer Prozentsatz der relevanten Gesamtpopulation befragt wurde, kann der Sampling Error als vernachlässigbar betrachtet werden.

Eine andere Fehlerquelle liegt im Coverage Error. Der Coverage Error tritt auf, wenn die Liste, von der die Stichprobe gezogen wird, nicht die gesamte relevante Population umfasst und somit nicht allen Subjekten die gleiche (oder bekannte) Chance eingeräumt wird, gezogen zu werden.³⁵ Da auch die Hoppenstedt-Datenbank keine vollständige Sammlung aller Controller und technischen Leiter umfasst, kann der Coverage Error nicht vollständig ausgeschlossen werden. Allerdings existiert kein plausibler Grund, warum die Personen in der Hoppenstedt-Datenbank in Bezug auf den Untersuchungsgegenstand dieser Studie systematisch von der Gesamtpopulation abweichen sollten. Dementsprechend kann die Hoppenstedt-Datenbank als Zufallssample der Gesamtpopulation aufgefasst werden, von der wiederum ein Zufallssample gezogen wurde. Auf der Basis dieser Annahme erscheint der Coverage Error nicht als bedrohlich für die Validität der Ergebnisse.

³⁰ Für eine Erläuterung der externen Validität vgl. Seite 116.

³¹ Vgl. Die Kommission der Europäischen Gemeinschaften 2003, S. 39.

³² Für eine Grundlageneinführung in das Analysieren komplexer Daten vgl. beispielsweise Lee/Forthofer 2006.

³³ Vgl. Dillman 2000, S. 11.

³⁴ Vgl. Dillman 2000, S. 11.

³⁵ Vgl. Dillman 2000, S. 11.

3.2 Entwicklung der Messmodelle

3.2.1 Recherche von existierenden Items und Konstrukten

Ein kritisches Element für den wissenschaftlichen Erkenntnisfortschritt ist die Entwicklung neuer Messungen.³⁶ In der methodologischen Literatur wird allerdings betont, dass neue Konstrukte nur dann entwickelt werden sollten, wenn keine etablierten Konstrukte vorliegen.³⁷ Für die Entwicklung des Fragebogens wurde deshalb zunächst die Literatur nach bereits existierenden Konstrukten durchsucht. Dazu wurden drei Arten von Quellen herangezogen: (1) Konstruktensammlungen, (2) psychologische Testdatenbanken und (3) Fachartikel.

(1) Für verschiedene Forschungszweige innerhalb der Betriebswirtschaftslehre existieren thematische Sammlungen von Konstrukten, die in der Vergangenheit im entsprechenden Fachgebiet zum Einsatz kamen. In verschiedenen Sammlungen wurden alle Variablen analysiert, um passende Konstrukte für die vorliegende Studie zu identifizieren. Die betrachteten Werke umfassen die vier Bände des Marketing Scales Handbook,³⁸ die Summary and Analysis of Multi-Item Scales Used in Logistics Research,³⁹ die Konstruktensammlung von Kwok/Sharp⁴⁰ und das Management Accounting & Control Scales Handbook.⁴¹

(2) Darüber hinaus wurden verschiedene Datenbanken mit psychologischen Tests durchsucht.⁴² So beispielsweise die Internetdatenbank Health and Psychosocial Instruments (HaPI) mit mehr als 100.000 Einträgen,⁴³ das Mental Measurement Yearbook mit über 4.000 Einträgen⁴⁴ und ETS Test Link mit über 25.000 Einträgen.⁴⁵

³⁶ So betont beispielsweise Churchill: "A critical element in the evolution of a fundamental body of knowledge [...] is the development of better measures of variables.;" Churchill 1979, S. 64.

³⁷ Vgl. DeVellis 2002, S. 154.

³⁸ Vgl. Bruner/Hensel 1992, Bruner/Hensel 1996, Bruner II/James/Hensel 2001 und Bruner II/Hensel/James 2005.

³⁹ Vgl. Keller et alii 2002.

⁴⁰ Vgl. Kwok/Sharp 1998.

⁴¹ Vgl. Schäffer 2008.

⁴² Diese Vorgehensweise folgt einer Empfehlung von DeVellis 2002, S. 155.

⁴³ Vgl. Ovid Technologies 2007.

⁴⁴ Vgl. Buros Institute of Mental Measurements 2007.

⁴⁵ Vgl. Educational Testing Service 2007.

(3) Außerdem wurden verschiedene Onlinedatenbanken mit Fachartikeln durchsucht; darunter Ebsco, ABI Inform/Pro Quest, ScienceDirect, Springerlink, SSRN, IDEAS-REPEC, Citeseer und scholar.google.com. Außerdem wurden Quellen beschafft, die in bereits gefundenen Artikeln zitiert und als möglicherweise passend erachtet wurden.

Insgesamt ist festzustellen, dass für den zu untersuchenden Sachverhalt *keine* Konstrukte, d. h. Multi-Item-Fragebatterien, vorliegen. Allerdings wurden einzelne Fragen gefunden, die bei der Entwicklung neuer Konstrukte verwendet werden konnten.

3.2.2 Entwicklung neuer Skalen gemäß C-OAR-SE

Da in der bisher publizierten Forschung keine geeigneten Konstrukte für die vorliegende Untersuchung existieren,⁴⁶ müssen neue Konstrukte entwickelt werden. In der Literatur finden sich verschiedene Prozessmodelle für die Entwicklung neuer Konstrukte. Einen einflussreichen⁴⁷ Vorschlag unterbreitete im Jahr 1969 Likert. Dieser beinhaltet die Stufen Initial Survey Design, Questionnaire Development und Data Analysis.⁴⁸ In den darauf folgenden Jahrzehnten wurden Abwandlungen entwickelt⁴⁹ und neue Prozesse vorgeschlagen.⁵⁰ Im Fachgebiet Marketing wurde beispielsweise mehr als zwanzig Jahre lang⁵¹ bei der Entwicklung neuer Skalen die Prozedur von Churchill angewandt.⁵² Vor einigen Jahren wurde dieses etablierte Verfahren durch Rossiter erweitert.⁵³ Dessen Methode trägt die Bezeichnung C-OAR-SE, wobei das Akronym für „Construct definition, Object classification, Attribute classification, Rater identification, Scale formation, and Enumeration and Reporting“⁵⁴ steht. Diamantopoulos diskutiert und kritisiert einige Details von Rossiters Ansatz, bezeichnet die Entwicklung der C-OAR-SE-Prozedur jedoch insgesamt als großen Verdienst für die akademische (Marketing-)Gemeinschaft.⁵⁵

⁴⁶ Vgl. Abschnitt 3.2.1 auf der vorherigen Seite.

⁴⁷ Vgl. Hensley 1999, S. 346

⁴⁸ Vgl. Likert 1969.

⁴⁹ Vgl. beispielsweise Spector 1992.

⁵⁰ Vgl. beispielsweise Schwab 1980.

⁵¹ Vgl. Rossiter 2002, S. 305.

⁵² Vgl. Churchill 1979, S. 66.

⁵³ Vgl. Rossiter 2002, S. 306 f.

⁵⁴ Rossiter 2002, S. 305.

⁵⁵ Vgl. Diamantopoulos 2005, S. 8.

Die Vorgehensweise in diesem Forschungsprojekt folgt im Wesentlichen der C-OAR-SE-Prozedur, die in Abbildung 3.1 dargestellt wird.

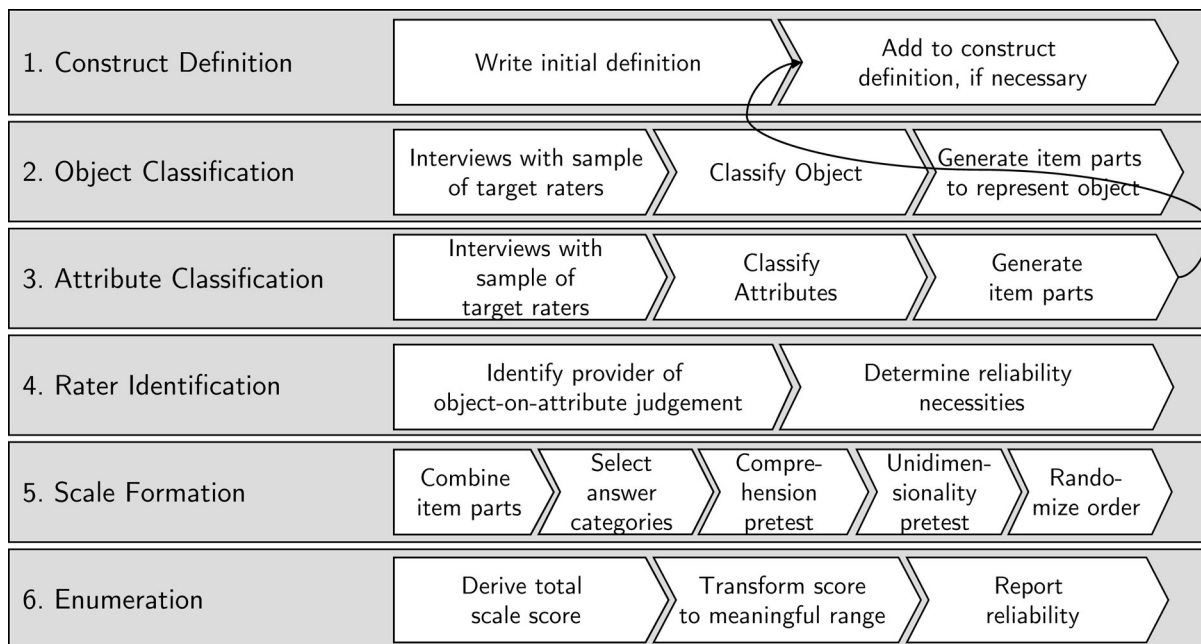


Abbildung 3.1: C-OAR-SE-Prozedur; in Anlehnung an Rossiter 2002, S. 306

In zwei Punkten wurde von der C-OAR-SE-Prozedur abgewichen. Zum einen wurde die Verwendung des O-A-R-Bewertungsformulars⁵⁶ nach zwei Expertentests eingestellt, da die Aufgabe von den Experten als sehr kompliziert angesehen wurde und die Ergebnisse so unterschiedlich waren, dass an der Praktikabilität und Reliabilität gezweifelt werden muss. Zum anderen konnte der von Rossiter geforderte Unidimensionalitätstest nicht mit einer konfirmatorischen Faktorenanalyse erfolgen, da die Zahl der Antworten im Pretest nicht den für diesen statistischen Test erforderlichen Umfang aufwies.⁵⁷

Die sechs Schritte der C-OAR-SE-Prozedur werden in den folgenden Abschnitten für die vorliegende Studie erläutert.

3.2.2.1 Definition und Erläuterung der Konstrukte

In einem ersten Schritt sind die Elemente des Modells genauer zu definieren, die in Kapitel 2.2 beschrieben und diskutiert wurden.⁵⁸ DeVellis betont in Bezug auf die Festlegung, was ein

⁵⁶ Vgl. Rossiter 2002, S. 321.

⁵⁷ Vgl. Abschnitt 3.4.1.8 auf Seite 91.

⁵⁸ Wie DeVellis pointiert feststellt: "Many researchers *think* they have a clear idea of what they wish to measure,

Konstrukt messen soll, insbesondere drei Anforderungen: eine Verankerung in der Theorie, eine möglichst starke Spezifizierung und eine Festlegung der Grenzen bzw. des Umfangs des Konstrukts.⁵⁹ Noch präziser sind die Forderungen von Rossiter. Die Definition sollte laut Rossiter das Objekt umfassen, seine Attribute und die Rater.⁶⁰ Unter Objekt wird dabei die Person, Firma oder Ähnliches verstanden, die bewertet wird. Das Attribut beschreibt die Dimension, nach der bewertet wird. Die Rater stellen die Personen dar, die geeignet sind, um Objekt und Attribut zu bewerten.

Tabelle 3.1: OAR-Definitionen von Projekteskalation und kognitiven Faktoren

Kurzbezeichnung	Objekt	Attribut	Rater
Projekteskalation	gescheitertes Projekt	ökonomisch-irrationales Festhalten an einem scheiternden Projekt	Projektkundige
Wahrnehmungsschwelle	Entscheidungsverantwortlicher	Grad der Schwierigkeit für den Verantwortlichen, negative Projektinformationen wahrzunehmen	Projektkundige
selektive Wahrnehmung	Entscheidungsverantwortlicher	Ausmaß, in dem der Entscheider bestimmte Informationen bevorzugt wahrnimmt	Projektkundige
Self-Justification-Druck	Entscheidungsverantwortlicher	Stärke der empfundenen Bedrohung von Image und Selbstbild bei einem Projektabbruch	Projektkundige
Sunk-Cost-Effekt	Entscheidungsverantwortlicher	ökonomisch-irrationale Tendenz, auf Grund bereits erfolgter, irreversibler Investitionen, Folgeinvestitionen in ein Projekt zu befürworten	Projektkundige
(übermäßiger) Optimismus	Entscheidungsverantwortlicher	(übermäßiger) Glaube an den Erfolg des Projekts	Projektkundige
Zeitpunkt des ersten Problemauftretens	Entscheidungsverantwortlicher	Einschätzung, wie spät im Projektverlauf die Probleme auftraten	Projektkundige
nicht hinterfragter Freiraum	Entscheidungsverantwortlicher	Freiraum, eigene Ziele ohne Hinterfragen in die Entscheidung über Abbruch oder Fortsetzung einfließen zu lassen	Projektkundige
Overconfidence	Entscheidungsverantwortlicher	Ausmaß des Vertrauens in die Zuverlässigkeit der Prognosen	Projektkundige

Projektkundige = als Projektkundige können Projektbeteiligte, aber auch Beobachter gelten, sofern sie ausreichende Kenntnisse haben, um das Projekt beurteilen zu können.

Die entwickelten Konstrukte entsprechen den genannten Anforderungen von DeVellis und Rossiter. Alle Konstrukte sind konzeptionell in Theorien verankert.⁶¹ Alle Konstrukte wurden spe-

only to find that their ideas are more vague than they thought.”; DeVellis 2002, S. 60; Hervorhebung im Original.

⁵⁹ Vgl. DeVellis 2002, S. 60 ff.

⁶⁰ Vgl. Rossiter 2002, S. 308 f.

⁶¹ Vgl. für eine detaillierte Erläuterung der theoretischen Fundierung Kapitel 2.2 auf Seite 21.

zifiziert und in ihrem Umfang geklärt: Eine detaillierte Auflistung der Object-Attribute-Rater- (OAR-) Definitionen bietet Tabelle 3.1 auf der vorherigen Seite. Um Redundanzen zu vermeiden wird für den Wortlaut der Definitionen der Konstrukte auf Kapitel 4.3 auf Seite 121 verwiesen.

3.2.2.2 Objekt-Klassifizierung

Der zweite Schritt, den Rossiter vorschlägt, besteht darin, Objekte als (1) konkret singular, (2) abstrakt kollektiv oder (3) abstrakt formativ zu klassifizieren.⁶²

(1) Als *konkret singular* wird ein Objekt dann bezeichnet, wenn alle Rater wissen, was das Objekt ist und es nur ein einzelnes Objekt für sie darstellt (beispielsweise eine bestimmte Person). Ebenso wird ein Objekt als konkret singular bezeichnet, wenn es zwar in verschiedenen Formen auftritt, aber an sich homogen ist (beispielsweise Coca-Cola in Flaschen, Dosen etc.)

(2) *Abstrakt kollektive* Objekte sind in den Augen der Rater heterogen, bilden für den Forscher jedoch eine Kategorie auf einem abstrakten Level. Unter dem Objekt kohlenstoffhaltige Getränke können beispielsweise Mineralwasser und Cola zusammengefasst werden. Für abstrakt kollektive Objekte stellt sich deshalb die Frage: Was beinhaltet das Objekt? Wird beispielsweise das Attribut „mögen“ verwendet, so ist es möglich, dass einige Rater Mineralwasser mögen, jedoch keine Cola und umgekehrt. Rossiter empfiehlt in einem solchen Fall, Untergruppen für das Objekt zu bilden und nach jedem einzeln zu fragen. Anschließend können die Bewertungen aggregiert werden, um einen Index für das abstrakt kollektive Objekt zu erhalten.

(3) Als *abstrakt formativ* werden Objekte klassifiziert, wenn verschiedene Rater unterschiedlicher Auffassung sind, aus welchen Komponenten sich das Objekt zusammensetzt (beispielsweise Kapitalismus). Für abstrakt formative Objekte stellt sich die Frage: Was bedeutet das Objekt? Die Komponenten eines abstrakt formativen Objekts müssen nicht unidimensional sein. Ein hohes Cronbach's α kann deshalb bei der Bewertung des Objekts kein Ziel sein.⁶³ Stattdessen empfiehlt Rossiter, Indizes zu bilden.⁶⁴

Im Zentrum der vorliegenden Studie stehen die beiden Objekte „abgebrochenes Projekt“ und „Entscheidungsverantwortlicher“.⁶⁵ Beide Objekte sind *konkret*, d. h. jeder Rater hat in etwa die

⁶² Vgl. Rossiter 2002, S. 309 ff.

⁶³ Vgl. Abschnitt 4.2.2.1 auf Seite 112.

⁶⁴ Vgl. Rossiter 2002, S. 312.

⁶⁵ Vgl. Tabelle 3.1 auf der vorherigen Seite.

gleiche Vorstellung von dem Objekt.⁶⁶ Die dritte Klasse von Objekten, die in der vorliegenden Studie untersucht werden, stellen die Maßnahmen des verhaltensorientierten Controllings dar. Auch hierbei handelt es sich um *konkrete* Objekte.

Die Elemente eines Items, die das jeweilige Objekt repräsentieren, lauten demnach schlicht „das Projekt“, „der Verantwortliche“ sowie die jeweilige Maßnahme.

3.2.2.3 Attribut-Klassifizierung

Der dritte Schritt der C-OAR-SE-Prozedur ist die Klassifizierung der Attribute, d. h. der Dimension auf der das Objekt bewertet wird. Rossiter unterscheidet dabei (1) concrete, (2) formed und (3) eliciting attributes.⁶⁷ Auf Grund der häufigeren Verwendung in der Literatur werden hier stattdessen die Bezeichnungen konkret, abstrakt formativ und abstrakt reflektiv verwendet.⁶⁸

(1) Als konkret werden Attribute bezeichnet, wenn sie direkt beobachtbare Sachverhalte bezeichnen und jeder Rater das Attribut identisch beschreiben würde. Bei abstrakten formativen und abstrakten reflektiven Attributen handelt es sich um nicht direkt beobachtbare, komplexe Phänomene. (2) Bei formativen Attributen setzt sich das Attribut aus verschiedenen Komponenten zusammen, die gemeinsam das Attribut bilden. (3) Reflektive Attribute hingegen entsprechen einem Zustand, der sich in verschiedenen äußeren Manifestationen widerspiegelt.⁶⁹

In der vorliegenden Studie weisen die kognitiven und moderierenden Faktoren reflektive Attribute auf. Die Einsatzhäufigkeit der Maßnahmen hingegen ist konkret. Die Attribute der Projekteskalation können als reflektiv zweiter Ordnung aufgefasst werden.⁷⁰ Ein solches Attribut besteht aus mehreren Komponenten (first-order), die selbst wiederum abstrakt (also formativ oder reflektiv) sind.⁷¹

⁶⁶ Damit ist nicht gemeint, dass alle Verantwortlichen gleich sind, sondern dass jeder Respondent eine konkrete Vorstellung davon hat, wer diese Person des Verantwortlichen ist. In diesem Zusammenhang ist daran zu erinnern, dass die Studie nur einzelne Verantwortliche betrachtet und keine Gruppenentscheidungen; vgl. 2.2.3 auf Seite 27.

⁶⁷ Vgl. Rossiter 2002, S. 313 ff.

⁶⁸ Vgl. beispielsweise Diamantopoulos/Winklhofer 2001, Eggert/Fassott 2005 und Temme 2006.

⁶⁹ Vgl. Rossiter 2002, S. 310.

⁷⁰ Vgl. hierzu die Ausführungen in Abschnitt 4.3 auf Seite 121.

⁷¹ Vgl. Rossiter 2002, S. 310.

Tabelle 3.2: Checkliste zur Formulierung von Items

Fragen für die Generierung und Überarbeitung von Items	Quelle
Enthält die Formulierung der Frage keinen Bias, der den Antwortenden in eine bestimmte Richtung lenken könnte? (Dies ist einer der häufigsten und schwerwiegendsten Fehler bei der Formulierung von Fragen.)	Vgl. Hague 1993, S. 64.
Verfügt der Befragte über das Wissen, um eine valide Auskunft zu geben?	Vgl. Hague 1993, S. 68.
Hat der Befragte die Antwort sofort parat?	Vgl. Dillman 2000, S. 34 ff.
Kann sich der Befragte an die vergangene Situationen ausreichend gut erinnern?	Vgl. Dillman 2000, S. 37.
Sind alle Fachbegriffe entfernt, die von den Befragten nicht verstanden werden könnten?	Vgl. Hague 1993, S. 65.
Sind die verwendeten Wörter möglichst einfach? Bei der Überarbeitung des Fragebogens wurden beispielsweise sophistizierte oder ungebräuchliche Wörter wie „salient“ ersetzt.	Vgl. DeVellis 2002, S. 67, Dillman 2000, S. 52, und Hague 1993, S. 66 f.
Enthalten die Fragen keine Abkürzungen, die den Befragten unbekannt sein könnten?	Vgl. Hague 1993, S. 65.
Werden mehrdeutige Wörter vermieden?	Vgl. DeVellis 2002, S. 68, und Hague 1993, S. 65.
Wie gut erfassen die Items die zentrale Idee des Konstrukts?	Vgl. DeVellis 2002, S. 66.
Wie klar und unmissverständlich ist die Formulierung?	Vgl. DeVellis 2002, S. 66.
Ist das Item so kurz wie möglich (ohne dass ein wesentlicher Bestandteil des Inhalts verlorengeht)?	Vgl. DeVellis 2002, S. 67, Dillman 2000, S. 53, und Hague 1993, S. 66.
Sind die Fragen so spezifisch wie möglich, d. h., werden vage Quantifizierungen vermieden, wenn präzisere Abfragen möglich sind?	Vgl. Dillman 2000, S. 54, und Hague 1993, S. 66.
Sind die Fragen möglichst ohne Verneinung formuliert?	Vgl. Hague 1993, S. 67.
Werden doppelte Verneinungen vermieden?	Vgl. DeVellis 2002, S. 68.
Enthalten alle Fragen nur genau einen Aspekt? Fragen, die mehrere unterschiedliche Aspekte enthalten, sind ein relativ häufiger Fehler bei Umfragen und werden im Englischen als „double-barreled item“ bezeichnet. Ein Beispiel für eine solche Frage, die im Verlauf der Überarbeitung aus diesem Fragebogen entfernt wurde, ist: „Wurden vom Controlling regelmäßige Berichte erstellt und vom Projektleiter auch gelesen?“ Diese Frage ist nicht zu beantworten, wenn zwar Berichte regelmäßig erstellt, aber meist nicht gelesen wurden.	Vgl. DeVellis 2002, S. 68.

Tabelle 3.2: Checkliste zur Formulierung von Items (Fortsetzung von Seite 57)

Fragen für die Generierung und Überarbeitung von Items	Quelle
Werden hypothetische Fragen nach Möglichkeit vermieden?	Vgl. Hague 1993, S. 67.
Sind die Befragten bereit, die geforderte Auskunft zu geben? Die unpersönliche Art der selbstständig zu bearbeitenden Online-Befragung ist hierbei ein Vorteil, denn es gibt empirische Befunde, dass Menschen bei selbstständig zu bearbeitenden Fragebögen ehrlicher antworten als bei persönlichen Interviews; vgl. hierzu die Diskussion und die Quellen in Abschnitt 3.1.1 auf Seite 45.	Vgl. Dillman 2000, S. 38.
Ist die Frage so formuliert, dass sie keinen der Befragten verärgern oder beleidigen kann?	Vgl. Hague 1993, S. 63.
Sind die Fragen im ganzen Satz gestellt?	Vgl. Dillman 2000, S. 54.
Sind die Fragen so formuliert, dass sie für alle Befragten eine Antwort erfordern? Dillman berichtet, dass Fragen die mit „Wenn ...“ beginnen (beispielsweise „Wenn Sie zum Essen ausgehen, bevorzugen Sie dann ...“), von einigen Befragten ignoriert werden, weil sie auf den Befragten nicht zutreffen (er also beispielsweise nie ausgeht).	Vgl. Dillman 2000, S. 34 f.

In den dritten Hauptschritt der Vorgehensweise nach C-OAR-SE fällt auch die Generierung der Items hinsichtlich der zu untersuchenden Attribute. Für den Wortlaut der generierten Items wird wiederum, um Redundanzen zu vermeiden, auf die Kapitel 4.3 auf Seite 121 und 4.5 auf Seite 169 verwiesen. Die Entwicklung der Attribute für jedes Item bildete in der vorliegenden Untersuchung den umfangreichsten Teil der Entwicklung der Items (da Objekt und Rater klar bestimmbar waren). Für das Generieren und Überarbeiten der Items existiert eine Reihe von Empfehlungen in der Literatur.

Tabelle 3.2 auf der vorherigen Seite fasst zusammen, welche Aspekte bei der Formulierung berücksichtigt wurden.

3.2.2.4 Rater-Klassifizierung

Der vierte Schritt besteht gemäß Rossiter in der Klassifizierung der Rater.⁷² Es werden die drei Kategorien (1) Individual Rater, (2) Expert Rater und (3) Group Rater unterschieden. (1) Ein Individual Rater trifft Aussagen über sich selbst. (2) Expert Rater stellen eine kleine Gruppe mit

⁷² Vgl. Rossiter 2002, S. 318 f.

Expertise im Hinblick auf das Konstrukt dar. (3) Als Group Rater werden beispielsweise Gruppen von Respondenten bezeichnet, die kein spezielles Expertenwissen aufweisen müssen. Dazu zählt Rossiter beispielsweise Kunden, Manager und Angestellte. Für die vorliegende Studie können die Projektkundigen als Group Rater klassifiziert werden. Mit projektkundig ist hierbei gemeint, dass die Respondenten ausreichende Kenntnisse über das Projekt aufweisen müssen. Dabei muss es sich jedoch nicht um Expertenwissen in Bezug auf die Konstrukte handeln (beispielsweise muss der Rater keine psychologischen Fachkenntnisse über kognitive Faktoren besitzen, um die Fragen der vorliegenden Studie sinnvoll beantworten zu können).

3.2.2.5 Skalenbildung

Die Bildung der Skalen beginnt mit der Kombination der zuvor generierten Item-Parts (Object, Attribut und Rater). Aus dem Zusammenfügen der Item-Parts entstehen die Indikatoren der zu messenden Variablen. Ebenso wie die Attribute können auch die Gesamtkonstrukte konkret, formativ oder reflektiv ausgeprägt sein.⁷³ Wenn alle drei Elemente konkret sind (Object, Attribut und Rater), liegt ein vollständig konkretes Konstrukt vor. Für die empirische Erhebung vollständig konkreter Konstrukte genügt ein einzelner Indikator. Für latente (formative und reflektive) Variablen werden Multi-Item-Konstrukte in der Forschung bevorzugt.⁷⁴

Ein Vorwurf in Bezug auf die Operationalisierung von Konstrukten besteht darin, dass formative Konstrukte häufig fälschlich als reflektive Konstrukte modelliert werden.⁷⁵ Jarvis/MacKenzie/Podsakoff geben Kriterien an, nach denen Konstrukte als reflektiv oder formativ klassifiziert werden können.⁷⁶ Homburg/Klarmann hingegen merken kritisch an, dass die kausale Richtung zwischen Indikator und latenter Variable häufig schwer zu bestimmen ist und die Entscheidung zwischen formativem und reflektivem Messmodell deshalb in der Regel stark subjektiv geprägt ist.⁷⁷ Auch wenn die Entscheidung zwischen formativer und reflektiver Messung subjektiv erfolgen muss, kritisiert Diller, dass diese Entscheidung in zahlreichen Studien gar nicht thema-

⁷³ Vgl. Abschnitt 3.2.2.3 auf Seite 56.

⁷⁴ Vgl. beispielsweise Bruner II/King 2003, S. 97 ff., für den Trend zu Multi-Item-Konstrukten in der Marketingforschung.

⁷⁵ Diese Kritik äußern beispielsweise Jarvis/MacKenzie/Podsakoff 2003, S. 206. Eine Studie von Fassott kommt für den deutschsprachigen Raum zu einem ähnlichen Ergebnis: Über 35 Prozent der untersuchten Konstrukte wurden als fehlspezifiziert eingestuft; vgl. Fassott 2006, S. 76.

⁷⁶ Vgl. Jarvis/MacKenzie/Podsakoff 2003, S. 203. Für eine Kritik an Jarvis/MacKenzie/Podsakoff vgl. Albers/Götz 2006, S. 13.

⁷⁷ Homburg/Klarmann 2006, S. 731.

tisiert wird.⁷⁸ Im Folgenden werden deshalb die Variablen der vorliegenden Studie kurz in die erläuterte Systematik eingeordnet.

Bei der Einsatzhäufigkeit der Maßnahmen handelt es sich in der vorliegenden Studie um ein *konkretes* Attribut. Daraus folgt, dass der Einsatz der Maßnahmen mit Single-Items abgefragt werden kann. Die kognitiven Faktoren (einschließlich der moderierenden Faktoren) weisen hingegen *abstrakte* Attribute auf, wie die Ausführungen zur Klassifizierung der Attribute gezeigt haben. Die Konstrukte, als Synthese aus Objekt, Attribut und Rater, sind demnach ebenfalls abstrakt. Entsprechend den methodischen Empfehlungen der Literatur werden diese Variablen deshalb als Konstrukte mit mehreren Items operationalisiert.⁷⁹

Für die Entscheidung zwischen formativen und reflektiven Konstrukten wurden die bereits erwähnten Regeln von Jarvis/MacKenzie/Podsakoff während der Entwicklung der Items⁸⁰ für jedes Konstrukt einzeln betrachtet.⁸¹ Als Ergebnis dieser Analyse zeigt sich, dass (1) für alle Konstrukte die Kausalrichtung vom Konstrukt zu den Items verläuft, (2) die Indikatoren des gleichen Konstrukts jeweils untereinander austauschbar sind, (3) davon auszugehen ist, dass die jeweiligen Indikatoren miteinander kovariieren und (4) die jeweiligen Indikatoren die gleichen Ursachen und Wirkungen aufweisen. Die Ausprägung aller Kriterien spricht somit für eine reflektive Operationalisierung der kognitiven Faktoren (einschließlich der moderierenden Faktoren).

Nachdem die Fragen gebildet wurden, müssen die Antwortmöglichkeiten festgelegt werden. Dabei ist eine Reihe von Aspekten zu beachten.⁸² Dazu gehört beispielsweise die Wahl der Anzahl der Skalenpunkte. Die Reliabilität steigt mit der Anzahl der Skalenpunkte: Bis fünf Skalenpunkte nimmt die Reliabilität stark zu, darüber hinaus steigt sie (nur) noch leicht an.⁸³ Einige Autoren sehen deshalb keinen Wert darin, mehr als fünf Skalenpunkte zu verwenden und stellen in Frage, dass Respondenten zwischen beispielsweise zehn Antwortoptionen sinnvoll differenzieren können.⁸⁴ Andere Autoren empfehlen ausdrücklich, mehr als fünf Antwortpunkte

⁷⁸ Vgl. Diller 2006, S. 613.

⁷⁹ So konstatiert DeVellis: "It is precisely this type of variable – one that is not directly observable and that involves thought on the part of the respondent – that is most appropriately assessed by means of a scale."; DeVellis 2002, S. 10.

⁸⁰ Vgl. die Erläuterungen ab Abschnitt 3.4.1.2 auf Seite 79 ff.

⁸¹ Vgl. Jarvis/MacKenzie/Podsakoff 2003, S. 203.

⁸² Vgl. beispielsweise Devlin/Dong/Brown 1993.

⁸³ Vgl. Lissitz/Green 1975.

⁸⁴ Vgl. Devlin/Dong/Brown 1993, S. 14. Insbesondere für telefonische Befragungen wird von langen Skalen abgeraten, da die Respondenten sonst überfordert sein könnten.

zu verwenden. So fordert beispielsweise Ping: “Specifically, to increase construct variance and reduce measurement error variance, the number of scale points could be increased by replacing the ubiquitous 5-point Likert scale with a 7-point Likert scale, a 10-point rating scale etc.”⁸⁵ Für die vorliegende Untersuchung werden siebenstufige Skalen gewählt, die einen angemessenen Kompromiss zwischen Differenzierungsmöglichkeit und Überschaubarkeit bieten.

Dabei wurde entsprechend den Empfehlungen aus der Literatur darauf geachtet, dass das End-of-Scale-Problem vermieden wird,⁸⁶ bei Bewertungsfragen die Mitte der Antwortskala dem neutralen Punkt entspricht,⁸⁷ der neutrale Punkt mit einer Null hervorgehoben wird⁸⁸ und die Kategorien der Antwortskala ausreichend beschriftet sind und nicht nur aus Ziffern bestehen.⁸⁹

Ebenfalls zur Skalenbildung gehören verschiedene Pretests. Diese werden in Kapitel 3.4.1 ausführlich erläutert.

3.2.2.6 Enumeration

Der letzte Schritt gemäß C-OAR-SE besteht in der Berechnung der Ergebnisse. Dazu gehört die Bestimmung der Durchschnitte für jedes Item einzeln sowie für die Gesamtkonstrukte. Diese Angaben sind in Abschnitt A.8 auf Seite 236 im Anhang aufgeführt. Eine ausführliche Darstellung der weiteren statistischen Prüfungen der Konstrukte befindet sich in Kapitel 4.3.

⁸⁵ Ping 2004, S. 134.

⁸⁶ Das End-of-Scale-Problem tritt auf, wenn durch ungünstig gewählte Ankerpunkte ein Großteil der Antworten am Ende der Skala liegt und dadurch die eigentlich vorhandenen Unterschiede nicht mehr gemessen werden können; vgl. Jamal 1993, S. 115.

⁸⁷ Unausgewogen wäre beispielsweise eine Skala, die aus den fünf Kategorien „unzufrieden“, „weder zufrieden noch unzufrieden“, „recht zufrieden“, „weitgehend zufrieden“ und „völlig zufrieden“ besteht, da die mittlere Kategorie („recht zufrieden“) nicht die neutrale Kategorie ist; vgl. Dillman 2000, S. 58.

⁸⁸ Hughes/Tippett/Thomas verwenden eine siebenstufige Likert-Skala, um die Zustimmung zu bewerten, wobei die Skala von „-3 (strongly disagree)“ bis „+3 (strongly agree)“ geht und bei „0“ einen neutralen Ankerpunkt aufweist. “This scale was used for two reasons: first, the positive and negative values more clearly connote a positive/negative sentiment of agreement, with ‘zero’ easily interpreted as neutral; second, the symmetry of scale reinforces the notion of balance in response options.”; Hughes/Tippett/Thomas 2004, S. 33.

⁸⁹ Vgl. Dillman 2000, S. 37.

3.3 Kontrolle potentieller Messfehler

Bereits in der Designphase des Fragebogens wurde versucht, potentielle Messfehler vorzusehen und deren Kontrolle zu ermöglichen. Denn wie Korman betont, “the point is not that adequate measurement is ‘nice’. It is necessary, crucial, etc. Without it we have nothing.”⁹⁰

3.3.1 Nonresponse-Bias

Das Problem einer abnehmenden Bereitschaft, Fragebögen auszufüllen, ist alles andere als neu.⁹¹ Kritisch diesbezüglich ist vor allem der Nonresponse-Bias, der dann entsteht, wenn die Teilnehmer einer Befragung in ihrer Einschätzung signifikant von den Nichtteilnehmern abweichen. Um den Nonresponse-Bias möglichst niedrig zu halten, sollte eine möglichst hohe Rücklaufquote angestrebt werden. Mit welchen Maßnahmen dies in der vorliegenden Untersuchung erfolgt ist, wird im Folgenden kurz dargestellt.

Da das Thema, bzw. der Bezug zu den persönlichen Interessen des Befragten, der wichtigste einzelne Einflussfaktor auf die Rücklaufquote ist,⁹² beginnt die Reduzierung des Nonresponse-Bias schon bei der Wahl des Titels der Untersuchung.⁹³ Die wichtigste Funktion des Titels besteht darin, die Befragten für das Thema zu interessieren, umso die Rücklaufquote zu steigern. Als Titel wurde deshalb gewählt: „Erfolgreiche Steuerung risikoreicher Projekte“.⁹⁴

Auch beim Anschreiben geht es vor allem darum, die Befragten zu motivieren, an der Befragung teilzunehmen. Dazu ist es wichtig, Interesse für das Thema zu wecken, Vertrauen zu erzeugen und den Befragten zu vermitteln, dass ihre Meinung wichtig ist und benötigt wird. Außerdem ist es oft angebracht, schon an dieser Stelle den Nutzen für die Befragten zu erwähnen, beispielsweise den Erhalt von Incentives und den Ergebnissen der Studie.⁹⁵ Die wesentlichen

⁹⁰ Korman 1974, S. 194.

⁹¹ Vgl. Steeh 1981.

⁹² Vgl. Hague 1993, S. 106.

⁹³ Dies ist insbesondere bei postalischen und Online-Befragungen ein wichtiger Punkt, da in diesem Fall keine persönliche Interaktion die Befragten zur Teilnahme motivieren kann.

⁹⁴ Wobei die Formulierung „erfolgreiche Steuerung“ für das Controller-Sample durch „erfolgreiches Controlling“ ersetzt wurde.

⁹⁵ Vgl. Hague 1993, S. 108.

Elemente, die für das Anschreiben berücksichtigt wurden, fasst Tabelle 3.3 auf der nächsten Seite zusammen.⁹⁶

Tabelle 3.3: Elemente des Anschreibens

Rücklaufbeeinflussende Elemente des Anschreibens	Quelle
Personalisierte Anrede, da diese zu einer signifikanten Steigerung der Rücklaufquote führt.	Vgl. Cook/Heath/Thompson 2000.
Das Anliegen des Fragebogens.	Vgl. Hague 1993, S. 108.
Appell an die moralische Verpflichtung. Da die Bereitwilligkeit, an einer Umfrage teilzunehmen, mit dem Gefühl der moralischen Verpflichtung steigt, wurde auf der Startseite die Bemerkung eingefügt: „Von Ihrer Teilnahme hängt der Erfolg dieses Forschungsprojekts ab! Nur durch die Auskünfte von Experten wie Ihnen können wir praxisbezogene Erkenntnisse gewinnen, die auch Ihnen zugute kommen.“	Vgl. Bosnjak/Tuten/Wittmann 2005, S. 501.
Eine Erläuterung, wie der Befragte ausgewählt worden ist.	Vgl. Hague 1993, S. 108.
Der Nutzen für den Befragten (Incentives etc.).	Vgl. Hague 1993, S. 108.
Klare Instruktionen, was zu tun ist und wie der Fragebogen zurückgesendet werden soll.	Vgl. Hague 1993, S. 108.
Eine Versicherung, dass das Ausfüllen einfach ist.	Vgl. Hague 1993, S. 108.
Eine Versicherung, dass alle Angaben vertraulich behandelt werden.	Vgl. Hague 1993, S. 108.
Dank an den Befragten.	Vgl. Hague 1993, S. 108.

3.3.2 Hindsight-Bias

Bei der Einschätzung vergangener Ereignisse unterliegen Menschen gelegentlich dem Hindsight-Bias (auch Outcome-Effekt genannt).⁹⁷ Dieser Effekt besteht darin, dass Individuen im Nachhinein der Ansicht sind, dass sie das eingetretene Ergebnis hätten voraussagen können. Im Nachhinein erscheint das Ergebnis offensichtlicher, als es in Wirklichkeit zuvor war.⁹⁸ Manch-

⁹⁶ Für den Wortlaut des Anschreibens vgl. Abschnitt A.4 auf Seite 220 im Anhang.

⁹⁷ Einer der ersten Nachweise erfolgte durch Fischhoff 1975. Für einen ausführlichen Literaturreview zu diesem Thema vgl. Hawkins/Hastie 1990.

⁹⁸ Brown/Solomon erläutern den Hindsight-Bias wie folgt: “Individuals with outcome knowledge will indicate that they could have predicted the reported outcome of an event with greater certainty than individuals without outcome knowledge will predict.”; Brown/Solomon 1993, S. 90.

mal wird dieser Effekt deshalb auch als „I knew it all along“-Effekt bezeichnet.⁹⁹ Ein Hindsight-Bias tritt beispielsweise dann auf, wenn Akteure nach ihrer Einstellung vor Eintreffen des Ereignisses gefragt werden, sich aber nicht mehr direkt an ihre damalige Einschätzung erinnern können.¹⁰⁰ Der Hindsight-Bias tritt *nicht* auf, wenn sich Personen an ihre damalige Einschätzung erinnern können.¹⁰¹ Dies ist beispielsweise dann der Fall, wenn ihre Einschätzung sehr einzigartig war oder eine hohe persönliche Bedeutung aufwies.¹⁰²

Es kann nicht mit Sicherheit ausgeschlossen werden, dass vor allem die beiden Konstrukte *Verspätung des Projektabbruchs* und *Bewertung des Ressourcen-Einsatzes*¹⁰³ vom Hindsight-Bias beeinflusst werden. Es sprechen jedoch zwei Argumente dafür, dass der Hindsight-Bias im vorliegenden Fall eher niedrig ist. (1) Eine geringere Wahrscheinlichkeit für den Hindsight-Bias liegt dann vor, wenn frühere Einschätzungen schriftlich festgehalten wurden, wie dies beispielsweise im Rahmen des Projektcontrolling der Fall sein kann.¹⁰⁴ (2) Ein weiteres Argument, das für eine geringere Ausprägung des Hindsight-Bias im Kontext der Eskalation des Commitments spricht, ist der Wettstreit um knappe Ressourcen (Budgets). Wenn Individuen um knappe Ressourcen wetteifern, werden die Erfolgchancen der einzelnen Projekte in der Regel genau beobachtet. Dies reduziert deutlich die Tendenz, im Nachhinein die ursprünglichen Einschätzungen zu vergessen.¹⁰⁵

Um die Ausprägung der genannten Aspekte beurteilen zu können, ist die Rolle der Respondenten im Projekt zu betrachten. Diese ist in der unteren Hälfte von Abbildung 4.2 auf Seite 109 dargestellt. 36 Prozent der Befragten hatten die Rolle des Projektleiters inne, 39 Prozent waren andere Projektbeteiligte im gleichen Unternehmen und 13 Prozent Projektsponsoren (Mehrfachnennungen möglich). Zumindest für Projektleiter unter den Respondenten war der Wettstreit um knappe Ressourcen mit hoher Wahrscheinlichkeit von großer Bedeutung. Aber auch für die anderen Projektbeteiligten sowie die Projektsponsoren ist davon auszugehen, dass die Projektressourcen eine erhebliche Bedeutung hatten. Dies spricht dafür, dass die Einstellung zum Projekt relativ stark bewusst war und eine nachträgliche Verzerrung deshalb geringer ausfällt.

⁹⁹ Vgl. für ein Experiment zum Hindsight-Bias beispielsweise Rudski 2002.

¹⁰⁰ Vgl. Jamal 1993, S. 114. Für eine Analyse der Erinnerungsinhalte im zeitlichen Verlauf vgl. Rieden 1959, S. 85 ff.

¹⁰¹ Vgl. Jamal 1993, S. 114.

¹⁰² Vgl. Hasher/Attig/Alba 1981.

¹⁰³ Vgl. Tabelle 4.5 auf Seite 124 und 4.6 auf Seite 125.

¹⁰⁴ Jamal führt diese Argumentation für ein Accounting System an; vgl. Jamal 1993, S. 114. Sie ist jedoch plausibel auf das Projektcontrolling übertragbar.

¹⁰⁵ Vgl. Jamal 1993, S. 114.

Insgesamt ist festzuhalten, dass der Hindsight-Bias zwar nicht ausgeschlossen werden kann, jedoch durch die Formulierung der Fragen minimiert wird und zusätzlich im gegebenen Kontext in eher geringer Ausprägung zu erwarten ist.

3.3.3 Informant-Bias

Wird bei einer Datenerhebung mittels einer Befragung nur eine Art von Informanten herangezogen, können deren Antworten systematisch durch ihre Position/Rolle verzerrt sein.¹⁰⁶ Dies wird als *Informant-Bias* bezeichnet.¹⁰⁷ „Das Ausmaß des Informant-Bias nimmt eine Größenordnung an, welche die Validität empirischer Befunde erheblich negativ beeinflusst. Im Durchschnitt ist der Methodenfehler (Informant-Bias) z. T. größer als die inhaltlich durch die Konstrukte erklärte Varianz.“¹⁰⁸

Ernst verlangt, dass in allen Studien (entgegen der häufigen Praxis beispielsweise in der Innovationsforschung) mindestens die Position der Befragten angegeben werden sollte. Dies erlaubt zumindest eine grobe Abschätzung des potentiellen Informant-Bias.¹⁰⁹ Dieser Forderung wird in der vorliegenden Untersuchung entsprochen.¹¹⁰ Zusätzlich greift die vorliegende Studie auf zwei verschiedene Arten von Informanten zurück: Entscheidungsverantwortliche und Nicht-Entscheidungsverantwortliche. Empirische Ergebnisse zeigen, dass Individuen, die verantwortlich für den Untersuchungsgegenstand sind (hier das abgebrochene Projekt), eine besonders hohe Tendenz zum Informant-Bias aufweisen.¹¹¹ Für die vorliegende Untersuchung ist dementsprechend besonders bei den Entscheidungsverantwortlichen von einem hohen Informant-Bias in Bezug auf die Projekteskalation und die kognitiven Faktoren auszugehen. Deshalb wurden im vorliegenden Forschungsprojekt zusätzlich Nicht-Entscheidungsverantwortliche befragt.¹¹² Diese Vorgehensweise ermöglicht es, die Ergebnisse der beiden Arten von Informanten ge-

¹⁰⁶ Vgl. beispielsweise Schermelleh-Engel/Schweizer 2007, S. 327 f.

¹⁰⁷ Für eine detaillierte Diskussion des Informant-Bias und seiner Bedeutung für die Managementforschung (am Beispiel der Neuproduktentwicklungsforschung) vgl. beispielsweise Ernst 2001, S. 87 ff. und 195 ff. Für eine empirische Untersuchung, die den Common-Source-Bias in der Leadership-Forschung nachweist, vgl. Dionne et alii 2002.

¹⁰⁸ Ernst 2003b, S. 1267.

¹⁰⁹ Vgl. Ernst 2003a, S. 23.

¹¹⁰ Vgl. Abschnitt 4.1.3 auf Seite 107.

¹¹¹ Vgl. Ernst 2003a, S. 23.

¹¹² Vgl. hierzu den Abschnitt 3.1.2 auf Seite 48, der die Auswahl der Informanten für die vorliegende Studie diskutiert.

genübertustellen und somit potentielle systematische Unterschiede zwischen verschiedenen Informanten aufzudecken. Die entsprechenden statistischen Analysen werden in Abschnitt 4.6.1 auf Seite 182 durchgeführt und ausgewertet. Noch genauer wäre dies mit einer dyadischen Erhebung möglich, bei der je ein Paar von Respondenten für das *gleiche* Projekt Auskunft gibt. Solche paarweisen Antworten sind jedoch erfahrungsgemäß nur schwer zu erhalten, so dass zugunsten einer höheren Rücklaufquote das geschilderte Erhebungsdesign gewählt wurde.

3.3.4 Common-Method-Biases

Als Common-Method-Variance wird die Varianz einer Messung bezeichnet, die durch die Messmethode anstatt durch das zu messende Konstrukt verursacht wird.¹¹³ Diese Varianz kann durch eine Reihe von Ursachen hervorgerufen werden, die als Common-Method-Biases oder Common-Source-Biases bezeichnet werden. Mögliche Ursachen sind beispielsweise „transient mood states, response styles, acquiescence, illusory correlations, social desirability, similarity of semantic content, and proximity to the criterion variable“.¹¹⁴

In der Literatur werden Common-Method-Biases seit über 40 Jahren diskutiert¹¹⁵ und von einigen Autoren als eine der Hauptfehlerquellen der empirischen Messung angesehen.¹¹⁶ Andere Autoren sind der Ansicht, dass das Problem der Common-Method-Variance überbewertet wird.¹¹⁷ Auch wenn die Verzerrungen durch den Common-Method-Bias häufig eher gering sind, sollten dennoch Maßnahmen ergriffen werden, um derartige Effekte ausschließen zu können.¹¹⁸ Im Folgenden soll anhand der Klassifizierung von Podsakoff et alii aufgezeigt werden, an welchen Stellen die Common-Method-Biases in der vorliegenden Arbeit berücksichtigt werden.

Podsakoff et alii unterteilen die Ursachen für Common-Method-Biases in vier Hauptgruppen mit insgesamt 21 Untergruppen: (1) Common Rater Effects, (2) Item Characteristic Effects, (3) Item Context Effects und (4) Measurement Context Effects.¹¹⁹

¹¹³ Vgl. Podsakoff et alii 2003, S. 879.

¹¹⁴ Lindell/Whitney 2001, S. 117.

¹¹⁵ Vgl. Campbell/Fiske 1959.

¹¹⁶ Vgl. Podsakoff et alii 2003, S. 879.

¹¹⁷ Vgl. Spector 2006, S. 221.

¹¹⁸ Vgl. Homburg/Klarmann 2006, S. 733.

¹¹⁹ Vgl. Podsakoff et alii 2003, S. 882.

(1) Common Rater Effects können entstehen, wenn abhängige und unabhängige Variablen vom gleichen Respondenten bewertet werden. Zum Common Rater Effect gehören beispielsweise die Untergruppen Social-Desirability-Bias, Leniency-Bias und Acquiescence-Bias. Diese Effekte und die eingesetzten Kontrollmaßnahmen werden in den Abschnitten 3.3.4.2, 3.3.4.3 und 3.3.4.4 behandelt.

(2) Item Characteristic Effects bezeichnen Effekte, die durch die spezifischen Eigenschaften eines Items hervorgerufen werden. Dazu zählen beispielsweise unklare Formulierungen und gemeinsame Antwortskalen bzw. Ankerpunkte. Diese Aspekte wurden in den Abschnitten 3.2.2.3 und 3.2.2.5 adressiert.

(3) Item Context Effects werden durch die spezifische Zusammenstellung der Items verursacht. Dazu gehört beispielsweise die Reihenfolge und Gruppierung der Items. Abschnitt 3.3.4.3 greift diese potentiellen Fehlerquellen auf.

(4) Measurement Context Effects beziehen sich auf die Begleitumstände der Befragung. Dazu gehören Ort, Zeit und gleiches Medium für die Befragung. Diese Aspekte wurden in Abschnitt 3.1.1 thematisiert.

Neben den zahlreichen Möglichkeiten, den Common-Method-Bias in seinen einzelnen Aspekten zu verringern und zu kontrollieren, kann auch versucht werden, den Gesamteffekt zu messen. Lindell/Whitney schlagen beispielsweise vor, ein oder mehrere Konstrukte in den Fragebogen aufzunehmen, die aus theoretischer Sicht mit mindestens einem anderen Konstrukt *nicht* korrelieren. Dies sollte insbesondere dann erfolgen, wenn die Befragung Korrelationen zwischen Konstrukten umfasst, die sich in der Vergangenheit besonders anfällig für die Method Variance (MV) gezeigt haben, wie beispielsweise Selbstauskünfte über Fähigkeiten und Leistungen.¹²⁰ “Marker-variable analysis [. . .] is far superior to overlooking MV effects altogether, which seems to be a very common way of addressing this problem.”¹²¹

Für den MV-Marker werden mehrere Empfehlungen abgegeben. (1) Der MV-Marker sollte möglichst zwischen den unabhängigen und den abhängigen Variablen positioniert werden. Damit kann der Nähe-Effekt (proximity) reduziert werden.¹²² (2) Der MV-Marker sollte bereits erprobt sein und nachweislich eine hohe Reliabilität aufweisen.¹²³ Das Konstrukt sollte also aus

¹²⁰ Vgl. Lindell/Whitney 2001, S. 119.

¹²¹ Lindell/Whitney 2001, S. 119.

¹²² Vgl. Lindell/Whitney 2001, S. 118.

¹²³ Vgl. zur Reliabilität die Erläuterungen in Abschnitt 4.2.2.1 auf Seite 112.

mehreren Items bestehen, für die bereits ein α empirisch ermittelt und dokumentiert wurde.¹²⁴ Da sich mit Hilfe des MV-Markers der Anteil der Common-Method-Variance in der statistischen Auswertung eliminieren lässt, der *für alle Variablen gleich* hoch ist, wird die Methode besonders wirkungsvoll, wenn die Items möglichst ähnlich in ihrer Anfälligkeit für Methodenvarianz sind.

Für die empirische Untersuchung der vorliegenden Arbeit wurde den Empfehlungen von Lindell/Whitney gefolgt. Als MV-Marker wurde das Konstrukt Kapitalmarktorientierung gewählt. Diese wurde im Fragebogen zwischen unabhängigen und abhängigen Variablen platziert und in Bezug auf die Antwortmöglichkeiten so gestaltet, wie die inhaltlichen Konstrukte. Das Konstrukt wird in Abschnitt 4.3.4.1 auf Seite 139 genauer erläutert und analysiert. Die Ergebnisse der Prüfung auf den Methodenbias mit Hilfe des MV-Markers werden in Abschnitt 4.6.2.1 erläutert.

Im Folgenden werden einige Common-Method-Biases näher diskutiert, die für die vorliegende Arbeit als besonders relevant erachtet werden.

3.3.4.1 Soziale Erwünschtheit und Impression-Management

Die bisherige Forschung deutet darauf hin, dass Menschen in Umfragen wahrheitsgemäß antworten, wenn keine Anreize zum Geben falscher Informationen bestehen.¹²⁵ Dabei ist jedoch zu berücksichtigen, dass auch nichtpekuniäre Anreize (beispielsweise gesellschaftliche Erwartungen) zu verzerrten Antworten führen können. Um dies zu vermeiden, sollten dem Individuum *ethisch neutrale* Fragen gestellt werden.¹²⁶ Dies wurde bei der Entwicklung der Konstrukte soweit wie möglich berücksichtigt.¹²⁷ Durch den Fokus der Arbeit auf kognitive Biases und Projektabbrüche lässt es sich jedoch nicht vermeiden, dass auch Aspekte angesprochen werden, die von dem einen oder anderen Befragten als unangenehm empfunden werden könnten.

Birnberg/Nath raten zu Vorsicht beim Interpretieren von Umfragen, wenn die Fragen heikle Themen betreffen.¹²⁸ Whyte formuliert es noch etwas zugespitzter: “When an individual is commanded by an organization to reveal his innermost feelings, he has a duty to himself to give

¹²⁴ Vgl. Lindell/Whitney 2001, S. 119.

¹²⁵ Unter Umständen antworten sie sogar wahrheitsgemäß, obwohl Anreize zum Geben falscher Informationen bestehen; vgl. Bohm 1972 und Smith 1979.

¹²⁶ Vgl. McFadden 1999, S. 96.

¹²⁷ Vgl. Abschnitt 3.2.2 auf Seite 52.

¹²⁸ Vgl. Birnberg/Nath 1967, S. 473.

answers that serve his self-interest rather than that of the Organization. In a word, he should cheat.”¹²⁹

Eine spezielle Ausprägung verzerrter Antworten ist die Tendenz, sozial erwünschte Auskünfte zu geben. Dies wird unter den Stichworten Social Desirability und Impression-Management diskutiert. Das Konzept der sozialen Erwünschtheit wurde 1957 von Edwards in die Diskussion über Messfehler eingebracht und bezeichnet die Tendenz, sich selbst in sozial erwünschter Weise zu beschreiben.¹³⁰ In den Folgejahren fand eine intensive Forschung zu dieser Problematik statt,¹³¹ in deren Zuge verschiedene Skalen entwickelt wurden, um die Tendenz einer Person zu messen, sozial erwünscht zu antworten.¹³²

Nach Einschätzung von Nunnally lassen die Forschungsergebnisse keinen Zweifel daran, dass viele psychologische Fragebögen so stark vom Problem sozialer Erwünschtheit betroffen sind, dass sich die Frage stellt, ob sie weiter verwendet werden sollten.¹³³ Soziale Erwünschtheit beeinflusst demzufolge nachweisbar die Messung von Persönlichkeitsvariablen,¹³⁴ von Einstellungen¹³⁵ und die Auskünfte über das eigene Verhalten.¹³⁶

Dementsprechend hoch ist die Relevanz der sozialen Erwünschtheit bei Forschungsprojekten, die Daten aus Selbstauskünften verwenden.¹³⁷ Im Jahr 1977 empfahl die American Psychological Association in ihren Standards for Educational and Psychological Tests, für *jedes* neu entwickelte Konstrukt die Korrelation mit der Tendenz zu sozial erwünschten Antworten anzugeben.¹³⁸ Wird bei aktuellen Studien keine Prüfung der Tendenz, falsch zu antworten, durch-

¹²⁹ Whyte 1956, S. 196.

¹³⁰ Vgl. Edwards 1957.

¹³¹ Für eine ausführliche Darstellung der frühen Forschung zur sozialen Erwünschtheit vgl. Crowne/Marlowe 1964.

¹³² Für eine Diskussion der Unterschiede zwischen den Skalen vgl. Edwards/Edwards/Clark 1988.

¹³³ Vgl. Nunnally 1967, S. 480.

¹³⁴ Vgl. Mick 1996.

¹³⁵ Vgl. Fisher 1993.

¹³⁶ Vgl. Mensch/Kandel 1988.

¹³⁷ Für eine Auflistung von Studien in verschiedenen Forschungsgebieten wie Psychologie, Organizational Behavior, Marketing u. v. m. vgl. Fisher/Katz 2000, S. 106. Bruner II/Hensel/James nennen allein für den Bereich Marketing über dreißig relevante Publikationen, die dieses Konstrukt verwendet haben; vgl. Bruner II/Hensel/James 2005, S. 586 ff. Für eine Auflistung von Social-Desirability-Konstrukten und deren Wirkung in Studien in sechs führenden Marketing-Journals von 1980 bis 1997 vgl. King/Bruner 2000, S. 86 f.

¹³⁸ Vgl. American Psychological Association 1977.

geführt, wie das mit einer „Lügen-Skala“, der „Social Desirability Scale“ oder ähnlichen Konstrukten der Fall ist, wird dies in einigen Artikeln als Limitation angeführt.¹³⁹

Zur Kontrolle des Social-Desirability-Bias werden verschiedene Verfahren vorgeschlagen, von denen zwei für die vorliegende Untersuchung angewendet wurden.

(1) Zum einen wird auf die Impression-Management-Subscale von Paulhus zurückgegriffen.¹⁴⁰ Diese Vorgehensweise deckt sich mit derjenigen von Bernardi.¹⁴¹ Um die Skala vom anglo-amerikanischen auf den deutschen Sprachraum zu übertragen, wurde das Backtranslation-Verfahren angewendet.¹⁴² Zusätzlich wurde in mehreren Vorstudien die kulturelle Übertragbarkeit geprüft¹⁴³ und eine Kurzversion der Skala entwickelt. Die Ergebnisse sind in Abschnitt 4.3.4.2 auf Seite 140 beschrieben.

(2) Der zweite verwendete Ansatz bezieht sich auf einen Vorschlag Jos und betrifft die Verwendung von direkten und indirekten Fragestellungen. Jos Idee besteht darin, dass die Aussage über einen hypothetischen *anderen* Manager durch die SelbstEinschätzung mitbestimmt wird.¹⁴⁴ In anderen Worten: Es wird angenommen, dass die Auskunft über einen anderen typischen Manager offenbart, was der Verantwortliche über sich selbst denkt, während seine Selbstdarstellung möglicherweise durch die soziale Erwünschtheit positiv verzerrt wird.

Jos Vorschlag baut auf den Erkenntnissen Fishers auf. Dieser hat nachgewiesen, dass der Social-Desirability-Bias weniger stark zum Tragen kommt, wenn eine Frage indirekt formuliert wird.¹⁴⁵ Dass eine indirekte Frage über andere Personen als Indikator für das eigene Verhalten dienen kann, basiert auf der Überlegung, dass Individuen Fragen nach ungewissen Zuständen unbewusst in Richtung ihres eigenen Zustands beantworten.¹⁴⁶ Beispiel: Wenn jemand angibt, dass andere Leute Alkohol trinken und danach Auto fahren, kann diese Aussage unter bestimmten Umständen als Indikator dafür dienen, dass die Person selbst zu diesem Verhalten neigt.¹⁴⁷

¹³⁹ Vgl. beispielsweise Lueg/Finney 2007, S. 37.

¹⁴⁰ Vgl. Paulhus 1986 und Paulhus 1992.

¹⁴¹ Vgl. Bernardi 2006, S. 47.

¹⁴² Vgl. Abschnitt 3.4.1.1 auf Seite 76.

¹⁴³ Vgl. Middleton/Jones 2000, S. 156.

¹⁴⁴ Vgl. Jo 2000.

¹⁴⁵ Vgl. Fisher 1993, S. 307 ff.

¹⁴⁶ Vgl. hierzu Campbell 1950, Holmes 1968 und Sherwood 1981. Ein experimenteller Nachweis für diese Annahme findet sich bei Fisher 1993, der wesentlich höhere Scores für indirekte Formulierungen sensibler Fragestellungen findet als für direkte.

¹⁴⁷ Vgl. Fisher 1993, S. 304.

Indirektes Fragen erlaubt den Befragten also, hinter der Fassade der Unpersönlichkeit ihre eigene Meinung auszudrücken.¹⁴⁸

Dem Ansatz von Jo folgend wird das Konstrukt zur selektiven Wahrnehmung einmal direkt und einmal indirekt abgefragt. Unter „indirekt“ wird hier verstanden, dass die gleiche Frage, die für den Entscheidungsverantwortlichen gestellt wurde, noch einmal für „einen anderen, durchschnittlichen Manager“ gestellt wird. Die entsprechenden Formulierungen sind in den Tabellen 4.10 auf Seite 132 und 4.21 auf Seite 142 nachzulesen. Dieses Konstrukt selektive Wahrnehmung wurde ausgewählt, da es am anfälligsten für soziale Erwünschtheit scheint; dies folgt dem Vorschlag Jos, weniger anfällige Konstrukte *nur* direkt und lediglich die gefährdetsten Konstrukte zusätzlich indirekt abzufragen.¹⁴⁹ Alle Konstrukte sowohl direkt als auch indirekt zu erheben, hätte den Fragebogen übermäßig verlängert.

Tabelle 3.4: Maßnahmen zur Verringerung des Impression-Managements

Maßnahme	Quelle
Die Teilnehmer wurden mehrfach darauf hingewiesen, dass ihre Angaben absolut vertraulich behandelt werden. Dies kann gleichzeitig dazu beitragen, die Rücklaufquote zu erhöhen.	Vgl. Kalton 1983, S. 65.
Die Teilnehmer wurden darauf hingewiesen, dass ihre Eingaben anonym, d. h. nur auf Branchenebene aggregiert, ausgewertet werden. Der Grad der Anonymität wird von verschiedenen Autoren als besonders wichtig erachtet, um Social-Desirability-Bias zu reduzieren.	Vgl. beispielsweise Fisher/Katz 2000, S. 117.
Neben den technischen Leitern wurden auch Controller befragt, die ihrer Position gemäß weniger Commitment als die direkt verantwortlichen Projektleiter aufweisen.	Vgl. hierzu die Ausführungen zum Informant-Bias in Abschnitt 3.3.3 auf Seite 65.
Die Formulierung der Fragen wurde sehr genau überdacht, um jede Art von Wertung auszuschließen.	Vgl. hierzu die Ausführungen in Tabelle 3.2 auf Seite 57.

Der Vorschlag Jos kann als Multitrait-Multimethod-Verfahren (MTMM) aufgefasst werden.¹⁵⁰ So werden Verfahren zur Prüfung der Konstruktvalidität bezeichnet, die aus einer Kombinati-

¹⁴⁸ Vgl. Simon/Simon 1975, S. 586.

¹⁴⁹ Vgl. Jo 2000, S. 140 f.

¹⁵⁰ Vgl. Campbell/Fiske 1959. Für Erläuterungen und Anwendungen vgl. beispielsweise Bagozzi/Yi 1991, Ernst 2003a, S. 12 f., Ernst 2003b und Kim/Lee 1997, S. 1272.

on von mehreren Merkmalen (traits) und mehreren Messmethoden (methods) bestehen.¹⁵¹ Dabei entsprechen in der vorliegenden Studie die Merkmale den verschiedenen Ausprägungen der Antwortmöglichkeiten bei den Variablen. Die verschiedenen Methoden sind (1) die Selbstauskunft zur selektiven Wahrnehmung durch den Entscheidungsverantwortlichen und (2) die Beurteilung der selektiven Wahrnehmung eines (hypothetischen) anderen, typischen Managers durch den Entscheidungsverantwortlichen.¹⁵²

Neben den beiden vorgestellten Verfahren zur Kontrolle des Impression-Managements wurde eine Reihe von Maßnahmen ergriffen, um den Effekt abzumildern. Diese Maßnahmen sind in Tabelle 3.4 auf der vorherigen Seite zusammengefasst.

3.3.4.2 Leniency-Bias

Eng verwandt mit dem Social-Desirability-Bias und dem Impression-Management ist der Leniency-Bias.¹⁵³ Dieser betrifft die Tendenz zu verzerrten Antworten in Richtung sozialer Erwünschtheit bei *Auskünften über andere Personen, mit denen man sich gut versteht*, also „the tendency to describe others in favorable but probably untrue terms“.¹⁵⁴

Die Kontrolle des Leniency-Bias kann analog zu den in Abschnitt 3.3.4.1 auf Seite 68 unter (1) und (2) vorgestellten Methoden erfolgen. Zusätzlich wird eine dritte Methode (3) mit Hilfe eines Freundschaftskonstrukts eingesetzt.

(1) Die im Rahmen dieser Arbeit entwickelte¹⁵⁵ Kurzversion von Paulhus' Impression-Management-Subscale¹⁵⁶ kann auch für die Prüfung des Leniency-Bias eingesetzt werden. Dazu wurden die Nicht-Entscheidungsverantwortlichen gebeten, die entsprechenden Fragen im Hinblick auf den Entscheidungsverantwortlichen zu beantworten. Die Fragen sind im Wortlaut in Tabelle 4.19 auf Seite 140 aufgeführt. Die verwendeten Items finden sich gleichzeitig in sehr ähnlicher Form in dem Leniency-Konstrukt von Schriesheim/Kinicki/Schriesheim,¹⁵⁷ so dass von einer hohen Validität auszugehen ist.

¹⁵¹ Vgl. Schermelleh-Engel/Schweizer 2007, S. 326.

¹⁵² Vgl. hierzu Schermelleh-Engel/Schweizer 2007, S. 327 und 332 ff.

¹⁵³ Dieser Begriff wurde von Guilford 1954 und Nunnally 1967 geprägt.

¹⁵⁴ Schriesheim/Kinicki/Schriesheim 1979, S. 1.

¹⁵⁵ Vgl. Abschnitt 3.4.1.1 auf Seite 76.

¹⁵⁶ Vgl. Paulhus 1986 und Paulhus 1992.

¹⁵⁷ Vgl. Schriesheim/Kinicki/Schriesheim 1979, S. 13.

(2) Auch zur Kontrolle des Leniency-Bias kann die oben diskutierte Methode von Jo beitragen.¹⁵⁸ Dazu wurden die Nicht-Entscheidungsverantwortlichen gegeben, die selektive Wahrnehmung nicht nur des Entscheidungsverantwortlichen, sondern zusätzlich auch die eines durchschnittlichen anderen Managers anzugeben. Die Fragen sind im Wortlaut in Tabelle 4.20 auf Seite 141 aufgeführt.

(3) Zusätzlich wurde ein Konstrukt für *Freundschaft* verwendet, um den Leniency-Bias zu prüfen. Denn wie Fisher experimentell nachgewiesen hat, sind Auskünfte über „den besten Freund“ genauso durch soziale Erwünschtheit beeinflusst wie Selbstauskünfte.¹⁵⁹ Der Wortlaut des Freundschaftskonstrukts ist in Tabelle 4.22 auf Seite 143 nachzulesen, die statistischen Auswertungen werden in Abschnitt 4.6.2.3 auf Seite 190 beschrieben.

3.3.4.3 Reihenfolge-Bias

Ein Reihenfolge-Bias liegt dann vor, wenn die Reihenfolge der Fragen Einfluss auf die Antworten hat. Dies kann beispielsweise durch die nachlassende Motivation der Teilnehmer verursacht werden. Darüber hinaus kann der Reihenfolge-Bias auch dadurch entstehen, dass die Befragten aus den gelesenen Fragen mentale Modelle über den Gesamtzusammenhang der Fragen aufstellen, die wiederum die Antworten auf die folgenden Fragen beeinflussen, da die Befragten versuchen, in ihrem mentalen Modell konsistent zu antworten.

In einigen Aspekten der Reihenfolge ist sich die Literatur weitgehend einig: Zweifellos am Anfang sollten Fragen stehen, die prüfen, ob die Person tatsächlich zum Ziel-Sample gehört und über die entsprechenden Kenntnisse zur Beantwortung des Fragebogens verfügt.¹⁶⁰ Einigkeit besteht ebenfalls über die Tatsache, dass die interessantesten Fragen so weit vorn wie möglich stehen sollten. Ein Fehler ist es deshalb, mit statistischen Fragen zur Person und zum Unternehmen zu beginnen. Neben dem Fokus auf spannende Fragen ist es auch wichtig, mit möglichst einfach zu beantwortenden Fragen zu beginnen, um den geschätzten Aufwand für den Befragten möglichst gering zu halten.¹⁶¹

Unbestritten ist auch, dass besonders schwierige oder heikle Fragen möglichst am Ende stehen sollten. Dafür sprechen zwei Argumente: (1) Zum einen ist die Abbruchwahrscheinlichkeit

¹⁵⁸ Vgl. Jo 2000, S. 145.

¹⁵⁹ Vgl. Fisher 1993, S. 311 f.

¹⁶⁰ Dies wurde für bei der durchgeführten Studie berücksichtigt, wobei die Antworten zu den Kenntnissen als Filtervariablen genutzt wurden, die den weiteren Verlauf des Fragebogens steuern. Für den Wortlaut der Fragen vgl. Abschnitt A.5 auf Seite 222 im Anhang.

¹⁶¹ Vgl. Hague 1993, S. 45.

dann geringer.¹⁶² Dies liegt an dem im Verlauf des Beantwortens aufgebautem Vertrauen und der damit einhergehenden Einstellung des Befragten.¹⁶³ Außerdem ist ein Abbruch unwahrscheinlicher, wenn der Befragte (anhand des Fortschrittsbalkens) sieht, dass er schon fast fertig ist.¹⁶⁴ (2) Wenn doch ein Abbruch erfolgt, ist er aus der Forschungsperspektive gegen Ende des Fragebogens weniger folgenschwer, da durch das Abschicken der vorherigen Abschnitte ein großer Teil der Daten bereits erfasst wurde und potentiell ausgewertet werden kann.

Schwieriger gestaltet sich die Frage, wie und ob die Items gruppiert werden sollen. In Bezug auf ihre Anordnung können verschiedene Fälle unterschieden werden. (1) Die Items der Konstrukte werden nicht für jedes Konstrukt als Block abgefragt, sondern alle Items werden zufällig vermischt.¹⁶⁵ Als Argument dafür wird angeführt, dass die Antworten eine höhere Validität aufweisen, wenn jedes Item einzeln beurteilt wird und die Items nicht im Kontext anderer Items oder eines Frames beurteilt werden. (2) Eine andere Möglichkeit besteht darin, die Items in logischen Blöcken anzuordnen. Dies erleichtert das Ausfüllen des Fragebogens und wird darum von einigen Forschern bevorzugt.¹⁶⁶ In der vorliegenden Untersuchung wird versucht, mit einer Kombination aus beiden Ansätzen ein Optimum zu finden. Die Items werden blockweise abgefragt, weil dadurch der logische Fluss leichter nachvollziehbar ist. Außerdem ermöglicht diese Variante kürzere Formulierungen durch einleitende Satzteile. Beides führt zu einer schnelleren Beantwortung, was sich positiv auf die Rücklaufquote und damit auf die Zuverlässigkeit der Ergebnisse auswirkt. Gleichzeitig werden die Möglichkeiten der onlinebasierten Befragung genutzt und die Items innerhalb eines Blocks zufällig (d. h. für jeden Antwortenden in einer anderen zufälligen Reihenfolge) angeordnet. Dadurch können Reihenfolgeeffekte innerhalb der Blöcke ausgeschlossen werden. Um Reihenfolgeeffekte über den gesamten Fragebogen zu kontrollieren, wurden zwei verschiedene Versionen des Fragebogens an jeweils die Hälfte des Controller-Samples versendet. Diese Versionen unterschieden sich in der Reihenfolge der drei inhaltlich zentralen Seiten des Fragebogens.¹⁶⁷ Die Auswertung dieses Verfahrens ist in Abschnitt 4.6.3 auf Seite 198 beschrieben.

¹⁶² Vgl. beispielsweise Kalton 1983, S. 65.

¹⁶³ Vgl. Hague 1993, S. 45.

¹⁶⁴ Es liegt nahe, hier den Sunk-Cost-Effekt als eine Ursache zu betrachten (vgl. Abschnitt 2.3.4.1 auf Seite 38), der ebenfalls vom Fertigstellungsgrad beeinflusst wird.

¹⁶⁵ Diese Vorgehensweise wenden beispielsweise auch Smith/Keil/Depledge 2001, S. 223, an.

¹⁶⁶ Vgl. beispielsweise Hague 1993, S. 45, und Hensley 1999, S. 350.

¹⁶⁷ Als aktuelles Beispiel für die unterschiedliche Anordnung von Blöcken vgl. beispielsweise d'Astous/Boujbel 2007, S. 233, und Lueg/Finney 2007.

3.3.4.4 Acquiescence-Bias

Befragte haben die Tendenz, Fragen zuzustimmen – unabhängig vom Inhalt der Frage. Dieses Phänomen wird auch als Acquiescence-, Affirmation- oder Agreement-Bias bezeichnet. Um diesen Effekt zu neutralisieren, kann die Methode des Reversed Coding angewendet werden.¹⁶⁸ Dabei werden einige Items so formuliert, dass sie eine Zustimmung zum Konstrukt ausdrücken, andere so, dass sie eine Ablehnung formulieren. Einige Autoren sind zudem der Auffassung, dass mit Hilfe von Reversed Coding die Aufmerksamkeit des Antwortenden erhöht wird.¹⁶⁹ In Bezug auf Reversed Coding gibt es zwei widerstreitende Fraktionen.¹⁷⁰

Während die Befürworter auf die oben genannten Argumente verweisen, führen die Gegner an, dass Reverse Coding die Validität von Antworten nachweisbar verringert¹⁷¹ und zu systematischen Messfehlern von Konstrukten beiträgt.¹⁷² DeVellis warnt ausdrücklich, dass die entgegengerichteten Items in einem Konstrukt häufig schlecht funktionieren, was dadurch hervorgerufen sein könnte, dass einige Befragte die gegensätzliche Bedeutung des Items übersehen.¹⁷³ Eine Untersuchung der Studien mit Reversed Items kommt zu der Schlussfolgerung, dass Reversed Items oft geringer auf einen Faktor laden als die positiv formulierten Items des gleichen Faktors.¹⁷⁴ Einige Forscher empfehlen deshalb, das Reversed Coding zu vermeiden, da die Probleme von Reversed Items ihren Nutzen überwiegen.¹⁷⁵ Aufgrund der divergierenden Meinungen wird der bewusste Verzicht auf Reversed Items *nicht* als methodisch zweifelhaft bewertet. Im Gegenteil wird beispielsweise die Studie von Jackson et alii,¹⁷⁶ die explizit keine Reversed Items verwendet, sogar als “excellent example of ‘best practices’”¹⁷⁷ bezeichnet. Unter Abwägung der genannten Vor- und Nachteile wurde für die vorliegende Studie bei der Neuentwicklung von Skalen auf Reversed Coding verzichtet.¹⁷⁸

¹⁶⁸ Diese Vorgehensweise wenden im Kontext der Eskalation des Commitments beispielsweise auch Smith/Keil/Depledge 2001, S. 224 ff., und Davis/DeZoort/Kopp 2006, S. 24 an.

¹⁶⁹ Vgl. für die theoretische Begründung Spector 1992 und für Beispiele von empirischen Studien, bei denen dieses Verfahren so begründet wurde, Hensley 1999, S. 350.

¹⁷⁰ Vgl. Hensley 1999, S. 350.

¹⁷¹ Vgl. Schriesheim/Hill 1981.

¹⁷² Vgl. Jackson et alii 1993.

¹⁷³ Vgl. DeVellis 2002, S. 69 f.

¹⁷⁴ Vgl. Hinkin 1995, S. 972.

¹⁷⁵ Vgl. DeVellis 2002, S. 70.

¹⁷⁶ Vgl. Jackson et alii 1993.

¹⁷⁷ Hinkin 1995, S. 974.

¹⁷⁸ Im Fall des Impression-Management-Index (vgl. Abschnitt 3.4.1.1 auf der nächsten Seite), der im englischen Original bereits Reversed Items enthält, wurden diese beibehalten.

3.4 Durchführung der empirischen Erhebungen

3.4.1 Vorstudien und Pretests

Schriftliche Fragebögen sollten stets Pretests unterzogen werden, um die Eignung für die Zielgruppe zu prüfen. So betont beispielsweise Hague, dass Fragebögen zu oft versagen, weil die Ersteller es nicht geschafft haben, die Fragen aus dem Blickwinkel der Befragten zu formulieren.¹⁷⁹ Darüber hinaus können Pretests auch dazu beitragen, die Validität der Messinstrumente zu verbessern. In den folgenden Abschnitten werden die Vorstudien und Pretests beschrieben, die für die vorliegende Studie durchgeführt wurden.

3.4.1.1 Vorstudie zur Entwicklung eines Impression-Management-Index

Wie in Abschnitt 3.3.4.1 auf Seite 69 erläutert, erscheint es sinnvoll, dem Fragebogen ein Kontrollkonstrukt zur Messung des Impression-Managements beizufügen. Da der existierende Index auf Englisch formuliert ist, musste eine deutschsprachige Version erstellt werden. Dazu wurden zunächst die Items der Impression-Management-Subscale von Paulhus¹⁸⁰ von mehreren Personen unabhängig voneinander ins Deutsche übersetzt und die Ergebnisse verglichen. Darauf basierend wurde eine deutsche Version erstellt, die anschließend von einem amerikanischen Muttersprachler zurückübersetzt wurde. Die neu erstellte englische Version der Fragen wurde mit den englischen Originalfragen verglichen. Dieses Verfahren wird als Backtranslation bezeichnet und kann helfen, Übersetzungsfehler aufzudecken. Es wurden keine nennenswerten Abweichungen entdeckt. Die Fragen sind im Wortlaut (englisch und deutsch) in Tabelle A.1 auf Seite 214 dargestellt.

Bei der Verwendung eines Impression-Management-Index müssen einige unerwünschte Nebenwirkungen bedacht werden. (1) Der Originalindex umfasste 20 Items.¹⁸¹ (2) Einige Fragen des ursprünglichen Index könnten von den Befragten als beleidigend aufgefasst werden. Damit verbunden ist die Gefahr sinkender Rücklaufquoten und fehlender Items.¹⁸² (3) Außerdem

¹⁷⁹ Vgl. Hague 1993, S. 41.

¹⁸⁰ Vgl. Bernardi 2006, Paulhus 1986 und Paulhus 1992.

¹⁸¹ Vgl. Paulhus 1986.

¹⁸² Vgl. Fisher 2000, S. 74.

stellt sich bei einigen Items die Frage, ob sie für die Zielgruppe sozial unerwünschtes Verhalten widerspiegeln oder stattdessen als irrelevant betrachtet werden.¹⁸³ Dieser Punkt ist umso relevanter, wenn der Index in anderen Kulturkreisen angewendet wird.¹⁸⁴

Den genannten Risiken wurde mit einer empirischen Vorstudie begegnet.¹⁸⁵ In einer Online-Befragung bewerteten 46 Personen aus dem deutschen Kulturkreis für jedes der 20 Items zwei Dimensionen: Zum einen wurde auf einer siebenstufigen Skala von „1 = völlig unwichtig“ bis „7 = extrem wichtig“ bewertet, wie wichtig im Sinne der sozialen Erwünschtheit jedes Item im deutschsprachigen Kulturkreis ist. Zum anderen wurde auf einer siebenstufigen Skala von „1 = gar nicht beleidigend“ bis „7 = extrem beleidigend“ abgefragt, wie beleidigend jede Frage im deutschsprachigen Kulturkreis wirkt.

Anschließend wurden fünf Items ausgewählt. Die Anzahl von fünf Items resultiert dabei aus der Einschätzung, wie viele Fragen mit Rücksicht auf die Gesamtlänge des Fragebogens möglich sind. Gleichzeitig wird die Anzahl der Fragen dadurch limitiert, dass die Impression-Management-Fragen für die Respondenten unangenehm sind und der Nutzen für sie nicht unmittelbar ersichtlich ist. Ob die Anzahl von fünf Items ausreicht, die Tendenz zum Impression Management reliabel zu messen, sollte mit Replikationen in zukünftigen Studien kritisch geprüft werden.

Für die Auswahl der Items wurde zunächst als notwendiges Kriterium darauf geachtet, dass die Relevanz im deutschsprachigen Kulturkreis hoch ist. Für die besonders relevanten Items wurde der Quotient aus Relevanz und beleidigender Wirkung gebildet. Höhere Werte entsprechen dabei Items, die für die Zielgruppe von Controllern und technischen Leitern relevant sind und gleichzeitig eine verhältnismäßig geringe beleidigende Wirkung aufweisen. Für die Auswahl wurden Items mit einem möglichst hohen Quotienten bevorzugt. Damit folgt die Vorgehensweise der Empfehlung Fishers, neue Konstrukte zu entwickeln, die „relatively short, inoffensive, and specific to the targeted attitude or behavior“¹⁸⁶ sind.

¹⁸³ Vgl. Fisher 2000, S. 75.

¹⁸⁴ Vgl. Middleton/Jones 2000, S. 149.

¹⁸⁵ Der Einleitungstext zu dieser Umfrage ist im Anhang in Abschnitt A.1 auf Seite 213 aufgeführt.

¹⁸⁶ Fisher 2000, S. 75.

Tabelle 3.5: Ergebnisse der Vorstudie zum Impression-Management

Nr.	Deutsche Übersetzung der Indikatoren	R	B	A
1.	Manchmal erzähle ich Lügen, wenn ich muss.	4,9	3,8	ja
2.	Ich vertusche meine Fehler nie.	4,4	3,1	ja
3.	Es gab Gelegenheiten, bei denen ich jemanden ausgenutzt habe.	5,0	4,3	
4.	Ich fluche nie.	2,6	1,8	
5.	Manchmal versuche ich etwas heimzuzahlen, statt zu vergeben und zu vergessen.	4,0	3,3	
6.	Ich befolge Gesetze immer, auch wenn es unwahrscheinlich ist, dass ich erwischt werde.	3,8	2,9	
7.	Ich habe schon Schlechtes über einen Freund hinter seinem Rücken gesagt.	5,2	4,2	ja
8.	Wenn ich höre, wie Leute sich privat unterhalten, versuche ich, nicht hinzuhören.	3,2	2,7	
9.	Ich habe schon zu viel Wechselgeld von einem Verkäufer erhalten, ohne es zu sagen.	2,8	2,5	
10.	Ich verzolle immer alles, was ich verzollen muss.	2,6	2,1	
11.	Als ich jung war, habe ich manchmal gestohlen.	4,8	4,5	
12.	Ich habe noch nie Abfall auf die Straße geworfen.	2,3	2,0	
13.	Ich fahre manchmal schneller als erlaubt.	1,8	1,8	
14.	Ich lese niemals erotische Bücher oder Magazine.	3,7	3,8	
15.	Ich habe schon Dinge getan, die ich niemandem erzähle.	3,5	3,2	
16.	Ich nehme nie Dinge, die mir nicht gehören.	5,2	4,0	ja
17.	Ich habe mich auf der Arbeit oder in der Schule schon einmal krank gemeldet, obwohl ich nicht wirklich krank war.	4,7	3,8	
18.	Ich habe noch nie ein Buch in der Bücherei oder einen Handelsartikel beschädigt, ohne es zu sagen.	3,1	2,8	
19.	Ich habe einige ziemlich schreckliche Angewohnheiten.	3,3	3,2	
20.	Ich tratsche nicht über anderer Leute Angelegenheiten.	4,2	3,2	ja

Bei den Zahlenangaben handelt es sich um den arithmetischen Mittelwert von 46 Respondenten.

R = Relevanz im deutschsprachigen Kulturkreis auf einer siebenstufigen Skala von „1 = völlig unwichtig“ bis „7 = extrem wichtig“.

B = Beleidigende Wirkung auf einer siebenstufigen Skala von „1 = gar nicht beleidigend“ bis „7 = extrem beleidigend“.

A = Ausgewählt für den verkürzten Impression-Management-Index.

Tabelle 3.5 fasst die Fragen und die Ergebnisse der Vorstudie zusammen. In der rechten Spalte ist gekennzeichnet, welche fünf Fragen für den Impression-Management-Index ausgewählt wor-

den sind. In Tabelle A.1 auf Seite 214 im Anhang sind die Indikatoren von Paulhus' Impression-Management-Subscale im englischsprachigen Original aufgeführt.¹⁸⁷ Die Ergebnisse der statistischen Prüfung des Impression-Management-Bias werden in Abschnitt 4.6.2.2 auf Seite 187 diskutiert.

3.4.1.2 Pretest 1: Zuordnung der reflektiven Items

Die im Folgenden beschriebenen Pretests 1 und 2 beziehen sich auf die Entwicklung der latenten Konstrukte für die Messung der kognitiven Faktoren. Die anschließenden Pretests 3 bis 7 beziehen sich auf den gesamten Fragebogen.

Eine Möglichkeit zu überprüfen, ob die Items den Definitionen der Konstrukte entsprechen, besteht darin, die Items unsortiert zu präsentieren und dann von Testern den einzelnen Konstrukten zuordnen zu lassen.¹⁸⁸ Hinkin hält einen solchen Sortierungsvorgang für *notwendig*, um die inhaltliche Validität (Content Validity) sicherzustellen.¹⁸⁹ Als Best Practices werden die Zuordnungstests von MacKenzie/Podsakoff/Fetter¹⁹⁰ und Butler¹⁹¹ genannt.¹⁹² Für diese Sortierung werden keine Experten benötigt, da es sich eher um eine kognitive Leistung handelt als um einen Abruf von Erfahrungswissen.¹⁹³ Hierbei bietet es sich an, als Kontrolle der Ernsthaftigkeit auch Items einzufügen, die klar in keine der Gruppen passen. Die vorliegende Arbeit folgt diesen Best-Practice-Empfehlungen.

Der Sortierungspretest umfasst neben den Variablen, die in der Hauptbefragung verwendet wurden,¹⁹⁴ noch drei weitere Variablen, da zum Zeitpunkt des ersten Pretests noch die Option offengehalten werden sollte, diese mit in die Hauptbefragung aufzunehmen. Die drei Variablen lehnen sich an die oberste Ebene unterhalb der Eskalationstendenz in Zayers Modell an¹⁹⁵ und werden im Folgenden als Wahrnehmungsdefizit, individuelles Zielsystem (in Anlehnung an die

¹⁸⁷ Paulhus 1986, zitiert nach Bernardi 2006, S. 51 f.

¹⁸⁸ Vgl. für diese Vorgehensweise beispielsweise Saraph/Benson/Schroeder 1989 und für eine Metaanalyse dieser und anderer Methoden Hardesty/Bearden 2004.

¹⁸⁹ Vgl. Hinkin 1995, S. 982.

¹⁹⁰ Vgl. MacKenzie/Podsakoff/Fetter 1991.

¹⁹¹ Vgl. Butler 1991.

¹⁹² Vgl. Hinkin 1995, S. 970.

¹⁹³ Vgl. Schriesheim/Hinkin 1990.

¹⁹⁴ Vgl. Abbildung 2.3 auf Seite 28.

¹⁹⁵ Vgl. Abbildung 2.2 auf Seite 22.

Variable „Änderung des Zielsystems“) und rationale Bewertung (in Anlehnung an die Variable „Bewertungsfehler“) bezeichnet.

Als Rater für den Sortierungspretest wurden 18 Freunde und Kollegen per E-Mail angeschrieben und darum gebeten, die internetbasierte Befragung auszufüllen. Von diesen Personen wurde eine nicht erreicht. Von den 17 erreichten Personen beantworteten 15 den Fragebogen. Dies entspricht einer Rücklaufquote 88 Prozent.

Auf der Webseite wurden die Teilnehmer mit einem kurzen Text begrüßt und gebeten, sich die Definitionen für die Konstrukte durchzulesen.¹⁹⁶ Anschließend wurden den Teilnehmern die Indikatoren präsentiert. Dabei waren die Indikatoren in Blöcken angeordnet und wurden innerhalb der Blöcke für jeden Teilnehmer in individuell zufälliger Reihenfolge angezeigt.¹⁹⁷ Die Teilnehmer wurden gebeten, jeden Indikator genau einem Konstrukt zuzuordnen. Die Personen nahmen also unabhängig voneinander eine Sortierung der zufällig angeordneten Items zu den Konstrukt-Kategorien (zusätzlich einer Kategorie „Sonstige“) vor.¹⁹⁸

Hinkin bezeichnete es als verwunderlich, dass die Zuordnungsprozedur die am häufigsten vernachlässigte Validitätsmaßnahme ist,¹⁹⁹ obwohl sie nach seiner Einschätzung einen verhältnismäßig wenig aufwendigen Entwicklungsschritt darstellt. Nach dem Feedback aus dem hier durchgeführten Zuordnungspretest ergibt sich jedoch eine differenziertere Bewertung. Hinkins Einschätzung mag zutreffen, wenn die Items nur wenigen verschiedenen Konstrukten zuzuordnen sind und die Konstrukte sich sehr klar unterscheiden.

Unter den Teilnehmern dieses Pretests, bei dem neun Zuordnungskategorien zur Auswahl standen und sich einige Konstrukte inhaltlich sehr nahe waren (beispielsweise Wahrnehmungsschwelle und Aufmerksamkeitsdefizit) vertraten mehrere die Ansicht, dass es sich bei der Sortierungsaufgabe um eine geistig sehr anspruchsvolle Aufgabe handle. Von einigen wurde dies als motivierende Herausforderung, von anderen als mühsam empfunden. Bei vielen und verwandten Kategorien dürfte der Sortierungspretest deshalb eine äußerst geringe Rücklaufquote aufweisen, wenn man ihn mit Praktikern durchführt.

¹⁹⁶ Für den Wortlaut des Begrüßungstextes und der Definitionen vgl. Abschnitt A.2 auf Seite 216.

¹⁹⁷ Diese Vorgehensweise verhindert einen Reihenfolge-Effekt; vgl. Abschnitt 3.3.4.3 auf Seite 73.

¹⁹⁸ Diese Anzahl von Testern wird in diesem Schritt als ausreichend betrachtet; so ließen beispielsweise MacKenzie/Podsakoff/Fetter 1991 ihre Items von zehn Doktoranden und Mitgliedern der Fakultät sortieren.

¹⁹⁹ Vgl. Hinkin 1995, S. 982, mit Bezug auf Schriesheim et alii 1993.

Tabelle 3.6: Ergebnisse des Sortierungspretests

Nr.	Wortlaut des Indikators	[%]
Wahrnehmungsdefizit		
1 a	Dem Entscheider waren die Probleme des Projekts erst sehr spät nach ihrem ersten Auftreten klar.	67
1 b	Während des Projekts sind Probleme aufgetaucht, die der Entscheider eigentlich schon früher hätte bemerken müssen.	67
1 c	Der Entscheider hat die Probleme des Projekts zu wenig wahrgenommen.	47
1 d	Der Entscheider hat wichtigen Problemen erst zu spät Aufmerksamkeit geschenkt.	79
1 e	Dem Entscheider sind wichtige Probleme zu spät bewusst geworden.	64
1 f	Dem Entscheider sind Informationen, die ein Fehlschlagen des Projekts angedeutet haben, erst sehr spät aufgefallen.	79
1 g	Die konkreten Probleme des Projekts waren dem Entscheider nur unzureichend bekannt.	77
Wahrnehmungsschwelle		
2 a	Wichtige negative Informationen sind für den Entscheider in der Menge der Informationen untergegangen.	53
2 b	Wichtige negative Informationen waren für den Entscheider schwer zu bemerken.	40
2 c	Wichtige negative Informationen waren für den Entscheider schwer von den unwichtigen Informationen zu unterscheiden.	53
2 d	Wichtige negative Informationen waren nicht auffällig genug, um vom Entscheider bemerkt zu werden.	93
2 e	Wichtige negative Informationen sind nicht bis zum Entscheider vorgedrungen.	71
2 f	Wichtige negative Informationen hatten es schwer, die Wahrnehmungsschwelle des Entscheiders zu überwinden.	93
2 g	Wichtige negative Informationen waren schwer also solche zu identifizieren.	93
selektive Wahrnehmung		
3 a	Wenn es neue Informationen zum Projekt gab, hat der Entscheider nur das gesehen, was er sehen wollte.	87
3 b	Wenn neue Informationen zum Projekt auftraten, schenkte der Entscheider positiven Informationen deutlich größeres Interesse als negativen Informationen.	80
3 c	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider Informationen weniger Beachtung geschenkt, wenn sie seinen bisher vertretenen Vorstellungen widersprachen.	67
3 d	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider die Themen, mit denen er sich gut auskannte, stärker beachtet.	79
3 e	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider negative Informationen ignoriert.	93
3 f	Wenn neue Informationen zum Projekt auftraten, wandte der Entscheider nur bestimmten Informationen seine Aufmerksamkeit zu (d. h., unterlag selektiver Wahrnehmung).	93
3 g	Wenn neue Informationen zum Projekt auftraten, fokussierte sich die Aufmerksamkeit des Entscheiders auf seine „Lieblingsthemen“.	79

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 15 Respondenten.

[%] = Prozentsatz der richtigen Zuordnungen

Tabelle 3.6: Ergebnisse des Sortierungspretests (Fortsetzung von Seite 81)

Nr.	Wortlaut des Indikators	[%]
individuelles Zielsystem		
4 a	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider sich von der Sorge beeinflussen lassen, dass ein Projektabbruch negative Konsequenzen für seine Karriere bedeutet hätte.	53
4 b	Bei seinen früheren Entscheidungen, das Projekt fortzuführen hat der Entscheider das Projekt auch deshalb weitergeführt, weil er sonst schlecht dagestanden hätte.	29
4 c	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider Rücksicht auf mögliche negative persönliche Folgen für sich selbst genommen und konnte deshalb nicht die Unternehmensziele verfolgen.	73
4 d	Bei seinen früheren Entscheidungen hat der Entscheider das Projekt weitergeführt, weil er vor sich selbst einen Abbruch als persönliches Versagen betrachtet hätte.	43
4 e	Bei seinen früheren Entscheidungen, das Projekt fortzusetzen, hat der Entscheider seine individuellen Ziele verfolgt, auch wenn sie den Unternehmenszielen widersprachen.	100
4 f	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat sich der Entscheider von seinen persönlichen Motiven leiten lassen.	100
4 g	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider so entschieden, weil ein Abbruch als sein persönliches Versagen ausgelegt worden wäre.	29
Self-Justification-Druck		
5 a	Der Entscheider hat sich persönlich für das Gelingen des Projekts verantwortlich gefühlt.	47
5 b	Vor sich selbst hätte der Entscheider einen Projektabbruch als persönliches Versagen empfunden.	67
5 c	Der Entscheider hatte Sorge um sein Image, falls es zu einem Projektabbruch kommen würde.	60
5 d	Der Entscheider befürchtete, dass seine Vorgesetzten bei einem Projektabbruch einen schlechten Eindruck von ihm haben könnten.	71
5 e	Der Entscheider erwartete bei einem Projektabbruch negative Auswirkungen auf seine Karriere.	50
5 f	Der Entscheider stand unter hohem Druck, das Projekt erfolgreich abzuschließen.	79
5 g	Der Entscheider empfand den Erfolg des Projekts als eng verknüpft mit seinem eigenen, persönlichen Erfolg.	21
rationale Bewertung		
6 a	Rein ökonomisch betrachtet war die Bewertung korrekt.	87
6 b	Die Bewertung erfolgte ökonomisch rational.	100
6 c	Die Gründe für die Fortführung des Projekts waren ökonomisch plausibel.	93
6 d	Die Bewertung war betriebswirtschaftlich einwandfrei.	100
6 e	Die Bewertung des Projekts erfolgte nach streng betriebswirtschaftlichen Gesichtspunkten.	100
6 f	In die Projektbewertung flossen nur ökonomische Kriterien ein.	100
6 g	Die Bewertung vergangener und zukünftiger Investitionen erfolgte unter rationalen Gesichtspunkten.	93

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 15 Respondenten.

[%] = Prozentsatz der richtigen Zuordnungen

Tabelle 3.6: Ergebnisse des Sortierungspretests (Fortsetzung von Seite 82)

Nr.	Wortlaut des Indikators	[%]
Sunk-Cost-Effekt		
7 a	Einen Projektabbruch sah der Entscheider wegen der bereits investierten unwiederbringlichen Investitionen (sunk costs) als Verschwendung von Ressourcen an.	80
7 b	Der Entscheider empfand den Wert des Projekts umso höher, je mehr Investitionen in das Projekt geflossen waren.	80
7 c	Der Entscheider sprach sich für eine Projektfortführung aus, weil bei einem Projektabbruch die zuvor erfolgten Investitionen umsonst gewesen wären.	100
7 d	Der Entscheider vertrat die Einstellung: „Wir haben schon zu viel investiert, um noch aufhören zu können.“	100
7 e	Je mehr Ressourcen unwiederbringlich in das Projekt investiert worden waren, desto „wertvoller“ erschien es dem Entscheider.	93
7 f	Ein Projektabbruch wäre als Vernichtung der investierten Ressourcen betrachtet worden.	93
7 g	Ein Abbruch des Projekts erschien zunehmend teurer, je mehr Ressourcen in das Projekt geflossen waren.	100
Optimismus		
8 a	Der Entscheider hat das Projekt nicht so riskant eingeschätzt, wie es war.	20
8 b	Die Bewertung des Projekts erfolgte lange Zeit zu positiv.	60
8 c	Die Probleme wurden in ihrer Bedeutung unterschätzt.	27
8 d	Der Entscheider zeigte sich lange Zeit optimistisch, was den erwarteten Ausgang des Projekts betraf.	100
8 e	Auch wenn es Probleme gab, war der Entscheider zuversichtlich, dass das Projekt am Ende erfolgreich sein würde.	93
8 f	Der Entscheider war überzeugt vom erfolgreichen Abschluss des Projekts.	100
8 g	Pessimismus war dem Entscheider fremd.	100

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 15 Respondenten.

[%] = Prozentsatz der richtigen Zuordnungen

Eine weitere Gefahr ist die sinkende Qualität der Angaben, wenn sich der Befragte demotiviert fühlt und deshalb die Fragen weniger aufmerksam beantwortet. Deshalb ist ein Sortierungspretest nicht uneingeschränkt zu empfehlen.²⁰⁰

²⁰⁰ Eine weitere Einschränkung liegt darin, dass viele Umfragen je Frageblock die gleichen Einleitungsworte für alle Indikatoren verwenden (vgl. beispielsweise die Frageblöcke im unteren Abschnitt von Seite 227). Diese gleichen Satzanfänge erleichtern den Respondenten das Ausfüllen. In einem Sortierungspretest ist jedoch nicht auszuschließen, dass das Schema der gleichen Satzanfänge von den Respondenten durchschaut und die Sortierung somit obsolet wird. Im vorliegenden Fall sprechen die Anzeichen jedoch gegen einen Einfluss der gleichen Satzanfänge während des Sortierungspretests. Denn zum einen wären dann weniger falsche Zuord-

Die Indikatoren und Ergebnisse des Sortierungspretests sind in den Tabelle 3.6 auf Seite 81 ff. dargestellt. Für die meisten Indikatoren liegt der Anteil der richtigen Zuordnungen über 65 Prozent. Eine genauere Untersuchung der Häufigkeiten (hier nicht abgebildet) zeigt, dass diejenigen Indikatoren, die häufig falsch zugeordnet wurden, oft demselben falschen Konstrukt zugeordnet wurden. So wurde beispielsweise der Indikator 2b in 40 Prozent der Fälle dem Konstrukt Wahrnehmungsschwelle richtig zugeordnet und in 33 Prozent der Fälle fälschlich dem Konstrukt Wahrnehmungsdefizit. Der Indikator 4b wurde von 29 Prozent der Respondenten richtig dem Konstrukt individuelles Zielsystem zugeordnet und von 71 Prozent fälschlich dem Konstrukt Self-Justification-Druck.

Dies zeigt, dass die wesentliche Ursache für falsche Zuordnungen nicht darin besteht, dass einzelne Indikatoren völlig unterschiedlich interpretiert wurden und zufällig über alle Konstrukte verteilt sind. Vielmehr besteht das Problem darin, dass einige Konstrukte konzeptionell zu eng beieinanderliegen und deshalb von den Respondenten nicht klar unterschieden werden können. Als Konsequenz wurden für den finalen Fragebogen die Konstrukte der ersten Ebene von Zayers Modell²⁰¹ weggelassen. Die Prüfung der Konstrukte Wahrnehmungsdefizit, individuelles Zielsystem und rationale Bewertung müssen demnach zukünftigen Studien vorbehalten bleiben. Die Vereinfachung des Modells wird in Abschnitt 2.2.2.3 auf Seite 26 detaillierter diskutiert.

3.4.1.3 Pretest 2: Bewertung der reflektiven Items

Besonders hohe Bedeutung für die Entwicklung *valider* Items wird in der Literatur Pretests mit Personen aus der Zielpopulation beigemessen.²⁰² Homburg/Klarmann fordern nachdrücklich, dass die endgültige Auswahl der Indikatoren stets unter Zuhilfenahme von Mitgliedern der Zielgruppe erfolgen sollte.²⁰³

nungen aufgetreten, zum anderen hat keiner der Teilnehmer im Freitextfeld, wo nach den angewendeten Sortierungsstrategien gefragt wurde, die Satzanfänge genannt.

²⁰¹ Vgl. Abbildung 2.2 auf Seite 22.

²⁰² Vgl. beispielsweise Ping 2004, S. 134.

²⁰³ Vgl. Homburg/Klarmann 2006, S. 732.

Tabelle 3.7: Ergebnisse des Bewertungspretests

Nr.	Wortlaut des Indikators	[%]
Wahrnehmungsdefizit		
1 a	Dem Entscheider waren die Probleme des Projekts erst sehr spät nach ihrem ersten Auftreten klar.	79
1 b	Während des Projekts sind Probleme aufgetaucht, die der Entscheider eigentlich schon früher hätte bemerken müssen.	88
1 c	Der Entscheider hat die Probleme des Projekts zu wenig wahrgenommen.	79
1 d	Der Entscheider hat wichtigen Problemen erst zu spät Aufmerksamkeit geschenkt.	88
1 e	Dem Entscheider sind wichtige Probleme zu spät bewusst geworden.	88
1 f	Dem Entscheider sind Informationen, die ein Fehlschlagen des Projekts angedeutet haben, erst sehr spät aufgefallen.	96
1 g	Die konkreten Probleme des Projekts waren dem Entscheider nur unzureichend bekannt.	79
Wahrnehmungsschwelle		
2 a	Wichtige negative Informationen sind für den Entscheider in der Menge der Informationen untergegangen.	87
2 b	Wichtige negative Informationen waren für den Entscheider schwer zu bemerken.	83
2 c	Wichtige negative Informationen waren für den Entscheider schwer von den unwichtigen Informationen zu unterscheiden.	88
2 d	Wichtige negative Informationen waren zu unauffällig, um vom Entscheider gut bemerkt zu werden.	87
2 e	Wichtige negative Informationen waren schwer zu bemerken, weil es keinen Anlass gab, sich mit ihnen auseinanderzusetzen.	61
2 f	Wichtige negative Informationen hatten es schwer, die Wahrnehmungsschwelle des Entscheiders zu überwinden.	74
2 g	Wichtige negative Informationen waren schwer also solche zu identifizieren.	91
selektive Wahrnehmung		
3 a	Wenn es neue Informationen zum Projekt gab, hat der Entscheider nur das gesehen, was er sehen wollte.	82
3 b	Wenn neue Informationen zum Projekt auftraten, schenkte der Entscheider positiven Informationen deutlich größeres Interesse als negativen Informationen.	91
3 c	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider Informationen weniger Beachtung geschenkt, wenn sie seinen bisher vertretenen Vorstellungen widersprachen.	91
3 d	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider die Themen, mit denen er sich gut auskannte, stärker beachtet.	91
3 e	Wenn neue Informationen zum Projekt auftraten, hat der Entscheider negative Informationen ignoriert.	55
3 f	Wenn neue Informationen zum Projekt auftraten, wandte der Entscheider nur bestimmten Informationen seine Aufmerksamkeit zu (d. h., unterlag selektiver Wahrnehmung).	96
3 g	Wenn neue Informationen zum Projekt auftraten, fokussierte sich die Aufmerksamkeit des Entscheiders auf seine „Lieblingsthemen“.	86

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 24 Respondenten.

[%] = Summe der Prozentsätze von „klar repräsentativ“ und „einigermaßen repräsentativ“.

Tabelle 3.7: Ergebnisse des Bewertungspretests (Fortsetzung von Seite 85)

Nr.	Wortlaut des Indikators	[%]
individuelles Zielsystem		
4 a	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider sich von der Sorge beeinflussen lassen, dass ein Projektabbruch negative Konsequenzen für seine Karriere bedeutet hätte.	95
4 b	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider das Projekt auch deshalb weitergeführt, weil er sonst schlecht dagestanden hätte.	91
4 c	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider Rücksicht auf mögliche negative persönliche Folgen für sich selbst genommen und konnte deshalb nicht die Unternehmensziele verfolgen.	91
4 d	Bei seinen früheren Entscheidungen hat der Entscheider das Projekt weitergeführt, weil er vor sich selbst einen Abbruch als persönliches Versagen betrachtet hätte.	83
4 e	Bei seinen früheren Entscheidungen, das Projekt fortzusetzen, hat der Entscheider seine individuellen Ziele verfolgt, auch wenn sie den Unternehmenszielen widersprachen.	83
4 f	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat sich der Entscheider von seinen persönlichen Motiven leiten lassen.	74
4 g	Bei seinen früheren Entscheidungen, das Projekt fortzuführen, hat der Entscheider so entschieden, weil ein Abbruch als sein persönliches Versagen ausgelegt worden wäre.	100
Self-Justification-Druck		
5 a	Der Entscheider hat sich persönlich für das Gelingen des Projekts verantwortlich gefühlt.	73
5 b	Vor sich selbst hätte der Entscheider einen Projektabbruch als persönliches Versagen empfunden.	77
5 c	Der Entscheider hatte Sorge um sein Image, falls es zu einem Projektabbruch kommen würde.	86
5 d	Der Entscheider befürchtete, dass seine Vorgesetzten bei einem Projektabbruch einen schlechten Eindruck von ihm haben könnten.	86
5 e	Der Entscheider erwartete bei einem Projektabbruch negative Auswirkungen auf seine Karriere.	77
5 f	Der Entscheider stand unter hohem Druck, das Projekt erfolgreich abzuschließen.	96
5 g	Der Entscheider empfand den Erfolg des Projekts als eng verknüpft mit seinem eigenen, persönlichen Erfolg.	86
rationale Bewertung		
6 a	Rein ökonomisch betrachtet war die Bewertung korrekt.	92
6 b	Die Bewertung erfolgte ökonomisch rational.	92
6 c	Die Gründe für die Fortführung des Projekts waren ökonomisch plausibel.	87
6 d	Die Bewertung war betriebswirtschaftlich einwandfrei.	83
6 e	Die Bewertung des Projekts erfolgte nach streng betriebswirtschaftlichen Gesichtspunkten.	92
6 f	In die Projektbewertung flossen nur ökonomische Kriterien ein.	83
6 g	Die Bewertung vergangener und zukünftiger Investitionen erfolgte unter rationalen Gesichtspunkten.	92

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 24 Respondenten.

[%] = Summe der Prozentsätze von „klar repräsentativ“ und „einigermaßen repräsentativ“.

Tabelle 3.7: Ergebnisse des Bewertungspretests (Fortsetzung von Seite 86)

Nr.	Wortlaut des Indikators	[%]
Sunk-Cost-Effekt		
7 a	Einen Projektabbruch sah der Entscheider wegen der bereits investierten unwiederbringlichen Investitionen (sunk costs) als Verschwendung von Ressourcen an.	96
7 b	Der Entscheider empfand den Wert des Projekts umso höher, je mehr Investitionen in das Projekt geflossen waren.	70
7 c	Der Entscheider sprach sich für eine Projektfortführung aus, weil bei einem Projektabbruch die zuvor erfolgten Investitionen umsonst gewesen wären.	100
7 d	Der Entscheider vertrat die Einstellung: „Wir haben schon zu viel investiert, um noch aufhören zu können.“	83
7 e	Je mehr Ressourcen unwiederbringlich in das Projekt investiert worden waren, desto „wertvoller“ erschien es dem Entscheider.	74
7 f	Ein Projektabbruch wäre als Vernichtung der investierten Ressourcen betrachtet worden.	83
7 g	Ein Abbruch des Projekts erschien zunehmend teurer, je mehr Ressourcen in das Projekt geflossen waren.	83
Optimismus		
8 a	Der Entscheider hat das Projekt nicht so riskant eingeschätzt, wie es war.	74
8 b	Die Bewertung des Projekts erfolgte lange Zeit zu positiv.	78
8 c	Die Probleme wurden in ihrer Bedeutung unterschätzt.	83
8 d	Der Entscheider zeigte sich lange Zeit optimistisch, was den erwarteten Ausgang des Projekts betraf.	96
8 e	Auch wenn es Probleme gab, war der Entscheider zuversichtlich, dass das Projekt am Ende erfolgreich sein würde.	96
8 f	Der Entscheider war überzeugt vom erfolgreichen Abschluss des Projekts.	91
8 g	Pessimismus war dem Entscheider fremd.	70

Bei den Zahlenangaben handelt es sich um den Anteil der gültigen Angaben von 24 Respondenten.

[%] = Summe der Prozentsätze von „klar repräsentativ“ und „einigermaßen repräsentativ“.

Eine Möglichkeit, dieser Forderung nachzukommen, besteht darin, die Zielgruppe bewerten zu lassen, wie gut die Items jeweils dem Konstrukt entsprechen.²⁰⁴ Diese Vorgehensweise wurde für die vorliegende Studie gewählt. Dazu wurden 36 Personen aus der Zielpopulation per E-Mail angeschrieben. Die Beteiligung an der Online-Befragung lag bei N = 24 (67 Prozent). In dem Fragebogen wurde jeweils eine Definition der Kernkonstrukte des Modells gegeben. Anschließend wurden die Teilnehmer gebeten, für jedes Item zu entscheiden, ob es „nicht repräsentativ

²⁰⁴ Für eine aktuelle Anwendung dieser Vorgehensweise vgl. beispielsweise d' Astous/Boujbel 2007, S. 233.

für das Konstrukt“, „einigermaßen repräsentativ für das Konstrukt“ oder „klar repräsentativ für das Konstrukt“ ist.

Die Indikatoren und Bewertungen für den zweiten Pretest sind der Tabelle 3.7 auf Seite 85 zu entnehmen. Die Fragen wurden seit dem ersten Pretest weiterer Überarbeitung unterzogen, so dass sich beispielsweise Indikator 2e vom ersten Pretest unterscheidet.

Der Bewertungspretest zeigt zufriedenstellende Ergebnisse. Alle Items wurden von der Mehrzahl der Befragten als einigermaßen oder klar repräsentativ für ihr jeweiliges Konstrukt eingeordnet.

Vergleicht man die Ergebnisse des Bewertungspretests (Pretest 2) mit dem Sortingpretest (Pretest 1), so fällt auf, dass einige Items beim Bewertungspretest deutlich besser abgeschnitten haben. Beispielsweise wurden die Items 8a und 8c im Sortingpretest nur von 20 und 27 Prozent der Befragten dem richtigen Konstrukt (Optimismus) zugeordnet. Im zweiten Pretest jedoch bewerteten 74 und 83 Prozent der Befragten diese beiden Indikatoren als einigermaßen oder klar repräsentativ für das Konstrukt Optimismus. Dies untermauert die zuvor geäußerte Hypothese,²⁰⁵ dass die falschen Zuordnungen im ersten Pretest vor allem dadurch verursacht wurden, dass die Konstrukte zu ähnlich sind, und nicht dadurch, dass die Indikatoren ihr Konstrukt nicht abbilden.

Für die späteren Kürzungen des Fragebogens dienten die Ergebnisse des Rating-Pretests als Anhaltspunkt, welche Items eine verhältnismäßig geringere Validität aufweisen und somit weggelassen werden können.

3.4.1.4 Pretest 3: Bewertung durch Wissenschaftler und Praktiker mit Survey-Erfahrung

Nachdem der erste Fragebogenentwurf vorlag, wurden nacheinander sechs Personen aus Wissenschaft und Praxis mit Expertise auf dem Gebiet großzahliger Umfragen gebeten, kritisches Feedback zu allen Fragen zu geben.²⁰⁶ Nach jeder Rückmeldung wurde der Fragebogen überarbeitet, bevor er an den nächsten Reviewer gesendet wurde. Im Ergebnis dieses Prozesses wurden vor allem Formulierungen überarbeitet und Fragen vereinfacht. Vereinzelt wurden Items gestrichen und neue aufgenommen.

²⁰⁵ Vgl. Abschnitt 3.4.1.2 auf Seite 79.

²⁰⁶ Vgl. Dillman 2000, S. 140 f.

3.4.1.5 Pretest 4: Zeitdauer

In der Vorbereitungsphase der empirischen Untersuchung wurden mehrfach Personen, die zuvor noch nicht mit dem Fragebogen in Berührung gekommen waren, gebeten, ihn auszufüllen und dabei die Zeit zu stoppen. Da manche Testpersonen für die Online-Version 35 Minuten benötigten statt der angestrebten 20 bis 25 Minuten, musste der Fragebogen gekürzt werden. Dazu wurde eine Reihe von Fragen entfernt, die nicht direkt mit dem Kern-Modell verknüpft waren, sondern sich auf den weiteren Kontext bezogen.

3.4.1.6 Pretest 5: Kognitive Interviews

Bei Umfragen besteht die Gefahr eines Messfehlers, wenn Fragen so ungünstig formuliert sind, dass sie ungenaue oder schwer interpretierbare Antworten zur Folge haben. Dies ist insbesondere bei selbstständig auszufüllenden Umfragen ein Problem, da der Antwortende keine Rückfragen stellen und kein sofortiges Feedback geben kann.²⁰⁷ Ziel des fünften Pretests war es deshalb, noch einmal explizit zu prüfen, ob alle Wörter verständlich sind und alle Fragen von jedem Respondenten beantwortet werden können. Für diesen Zweck bietet sich die Technik des kognitiven Interviews (cognitive interviewing) besonders an.²⁰⁸ Dabei wird ein Respondent gebeten, den Fragebogen auszufüllen und gleichzeitig „laut zu denken“. Der zuhörende Forscher kann mitvollziehen, welche spontanen Assoziationen und Gedanken jede Frage auslöst, wie sie interpretiert wird und bei welchen Fragen Probleme auftreten. Rossiter bezeichnet in seinem Artikel zur Konstruktentwicklung mittels der C-OAR-SE-Prozedur²⁰⁹ kognitive Interviews als die beste Methode zum Pretest von Skalenitems.²¹⁰ Das Verfahren wurde deshalb mit vier Personen der Zielgruppe durchgeführt.

Im Zuge dieses Pretests äußerten mehrere Befragte ihre Verwunderung darüber, dass sich einige Fragen auf den (scheinbar) gleichen Sachverhalt beziehen. Auch wenn sich dieser Eindruck bei reflektiven Messungen nicht immer ganz vermeiden lässt,²¹¹ kann die Irritation der Befragten zumindest reduziert werden, indem auf die Messung sehr ähnlicher Konstrukte oder Indikatoren verzichtet wird. Der Fragebogen wurde deshalb weiter gekürzt.²¹²

²⁰⁷ Vgl. Dillman 2000, S. 11.

²⁰⁸ Vgl. Bolton 1993 und Schwarz 1999.

²⁰⁹ Vgl. Kapitel 3.2.2 auf Seite 52.

²¹⁰ Vgl. Rossiter 2002, S. 320.

²¹¹ Vgl. Diller 2006, S. 615.

²¹² Vgl. hierzu Abschnitt 2.2.2.3 auf Seite 26.

3.4.1.7 Pretest 6: Funktionalität des Web-Interfaces

Bei der Programmierung des Online-Fragebogens wurden die Hinweise aus der Literatur berücksichtigt, um eine technisch bedingte Reduzierung der Rücklaufquote so weit wie möglich zu vermeiden.²¹³

Zudem wurde mit vier Personen aus der Praxis ein Funktionstest durchgeführt, der auch als *Piloting* bezeichnet wird. Wichtig dabei ist es, das gleiche Medium zu benutzen wie bei der geplanten Hauptbefragung.²¹⁴ Da die Einschätzung der Bedienbarkeit stark von den Fähigkeiten des Nutzers abhängt, wurden auch für diesen Pretest verschiedene Personen gebeten, den Fragebogen auszufüllen. Da Analysen verschiedener Umfragen gezeigt haben, dass die Item-Nonresponse mit dem Alter des Respondenten und dessen Ausbildung korreliert, empfehlen Craig/McCann, in die Pretests besonders auch ältere Teilnehmer einzubeziehen und unterschiedliche Bildungsgrade zu berücksichtigen.²¹⁵ Dieser Empfehlung wurde für die vorliegende Studie Folge geleistet. Der Online-Fragebogen wurde deshalb auch von einer Person im Alter von über 65 Jahren und von einer Person ohne Hochschulabschluss getestet.

Als Resultat dieses Pretests wurden einige zusätzliche Hinweise an Stellen eingefügt, an denen Missverständnisse aufgetreten waren. Beispielsweise wurde darauf hingewiesen, dass die Antwortmöglichkeiten eine „Skala“ darstellen, dass also nicht nur die belabelten Endpunkte, sondern auch alle Punkte dazwischen angekreuzt werden können. Darüber hinaus wurden auf der Startseite die Links zu Hintergrundinformationen so geändert, dass sie nicht mehr direkt anklickbar sind. Die direkten Links hatten dazu geführt, dass ein Tester sich auf den verlinkten Seiten „verirrte“ und nicht mehr auf die Folgeseiten des Fragebogens gelangte. Um ganz sicher zu gehen, dass jeder Respondent weiß, wie er zum Fragebogen gelangen kann, wurde zusätzlich der Hinweis eingefügt „Zum Start des Fragebogens klicken Sie bitte auf ‚Weiter‘.“

²¹³ Vgl. beispielsweise Fraley 2004, S. 287 ff., für die folgenden Hinweise: verschiedene Auflösungen testen, Scrollen minimieren, insb. in Bezug auf Anker von Skalen, klare Kontraste zwischen Schrift und Hintergrund, kein Java Script verwenden – da es nicht mit allen Browsern gleich funktioniert und elf Prozent der User Java Script deaktiviert haben, Test mit verschiedenen Browsern, keine Plug-Ins verwenden.

²¹⁴ Vgl. Hague 1993, S. 96.

²¹⁵ Vgl. Craig/McCann 1978, S. 289.

3.4.1.8 Pretest 7: Großzahlige Erhebung zur Prüfung der Unidimensionalität

Kritisiert werden in der Literatur Studien, die nur eine Erhebung durchführen und das gleiche Sample für die Konstruktentwicklung und den Hypothesentest verwenden.²¹⁶ Die Vorgehensweisen von Churchill und Rossiter beinhalten deshalb jeweils einen großzahligen Pretest und eine Hauptbefragung. Genau genommen dient diese Befragung jedoch auch zur Validierung der Skala. Das heißt, eine Verwendung zum Test von Hypothesen dürfte dann erst in einer dritten großzahligen Untersuchung erfolgen.²¹⁷ Gleichzeitig müssen Aufwand und Nutzen weiterer großzahliger Pretests gegeneinander abgewogen werden. Denn “construct validation is indeed a never-ending process”.²¹⁸

Tabelle 3.8: Ergebnisse des großzahligen Pretests

Skala	Anzahl Indikatoren	Cronbachs α
Verspätung des Projektabbruchs	3 Indikatoren	0,83
Beurteilung des Ressourceneinsatzes	3 Indikatoren	0,74
Investierter Anteil der geplanten Ressourcen	3 Indikatoren	0,85
Wahrnehmungsschwelle	5 Indikatoren	0,91
selektive Wahrnehmung	4 Indikatoren	0,93
Self-Justification-Druck	4 Indikatoren	0,90
Sunk-Cost-Effekt	4 Indikatoren	0,86
Optimismus	4 Indikatoren	0,87
Zeitpunkt des ersten Problemauftretens	3 Indikatoren	0,91
Nicht hinterfragter Freiraum	4 Indikatoren	0,92
Kapitalmarktorientierung	3 Indikatoren	0,94
selektive Wahrnehmung anderer Manager	4 Indikatoren	0,93
Freundschaft	3 Indikatoren	0,87

Die Angaben beziehen sich auf 84 Respondenten mit Ausnahme der Freundschafts-Skala, die nur die Nicht-Entscheidungsverantwortlichen (N = 44) betrifft.

Die Population der großzahligen Vorstudie muss der zukünftigen Zielgruppe, auf die generalisiert werden soll, entsprechen.²¹⁹ Über die Größe des Pretest-Samples finden sich in der Literatur verschiedene Angaben.²²⁰ Nunnally empfiehlt zehn Subjekte pro Item,²²¹ Spector

²¹⁶ So stellen Flynn/Pearcy die Frage: “What can be reasonably said about a scale developed in one study? Can it possibly have demonstrated reliability and validity to any real extend? It should not be surprising that when another researcher uses such a measurement tool, it does not work.”; Flynn/Pearcy 2001, S. 413.

²¹⁷ Vgl. Flynn/Pearcy 2001, S. 413.

²¹⁸ Anastasi 1986, S. 4.

²¹⁹ Vgl. Hinkin 1995, S. 971.

²²⁰ Für eine Diskussion verschiedener Empfehlungen vgl. DeVellis 2002, S. 136 f., und Hinkin 1995, S. 973.

²²¹ Vgl. Nunnally 1967, S. 257.

schlägt 100 bis 200 Subjekte vor.²²² Forschungen von Guadagnoli/Velicer haben ergeben, dass in den meisten Fällen eine Samplegröße von 150 Personen für *exploratorische* Faktoranalysen genügt.²²³ Für *konfirmatorische* Faktoranalysen werden mindestens 200 Datensätze empfohlen.²²⁴

Für den großzahligen Pretest wurden 1.110 Controller und technische Leiter angeschrieben. 871 E-Mails wurden erfolgreich zugestellt. Von den Antworten konnten 84 ausgewertet werden, was einer Rücklaufquote von 9,6 Prozent entspricht. In Tabelle 3.8 auf der vorherigen Seite wird Cronbachs α für jede Skala angegeben. Alle Werte liegen über dem geforderten Mindestwert von 0,7.²²⁵ Da die absolute Anzahl von 84 Antworten nicht der oben genannten Größenordnung für eine exploratorische oder konfirmatorische Faktorenanalyse entspricht, wurde keine Faktorenanalyse durchgeführt. Somit werden auf der Basis des Pretests auch keine Items eliminiert. Für eine detaillierte Analyse der Skalen einschließlich konfirmatorischer Faktorenanalyse wird deshalb auf Abschnitt 4.3 auf Seite 121 verwiesen.

Der großzahlige Pretest führte nicht zu Veränderungen im Fragebogen. Deshalb wurde nach der Durchführung der Hauptbefragung geprüft, ob die Datensätze des Pretests mit denen der Hauptbefragung zusammengelegt werden können. Dazu wurden Mittelwerte der Indikatoren der beiden Gruppen mittels t-Test verglichen. Zwischen dem großzahligen Pretest-Sample und der Hauptumfrage traten keine nennenswerten Unterschiede auf. Für die statistischen Analysen der folgenden Abschnitte werden deshalb beide Samples gemeinsam betrachtet.

3.4.2 Durchführung der Hauptbefragung

Die Hauptumfrage fand von Mai bis August 2007 statt. Die Adressaten²²⁶ wurden mit einer personalisierten E-Mail kontaktiert. Der Text des Anschreibens ist im Anhang in Abschnitt A.4 auf Seite 220 abgedruckt.

²²² Vgl. Spector 1992.

²²³ Vgl. Guadagnoli/Velicer 1988.

²²⁴ Vgl. Hoelter 1983.

²²⁵ Vgl. Pallant 2005, S. 90.

²²⁶ Zur Auswahl der Befragten vgl. Abschnitt 3.1.2 auf Seite 48.

Die Teilnehmer wurden zunächst gefragt, ob sie Erfahrungen mit abgebrochenen und/oder gescheiterten Projekten haben. Je nach ihrer Antwort wurden sie gebeten, sich an ein bestimmtes Projekt zu erinnern und zu diesem konkreten Projekt alle folgenden Aussagen zu treffen.²²⁷

Reminder, d. h. Erinnerungen, haben sich als eines der wirkungsvollsten Mittel herausgestellt, um die Rücklaufquote zu erhöhen.²²⁸ Im Rahmen der vorliegenden Untersuchung wurden deshalb drei Erinnerungsschreiben versendet. Die Reminder wurden so formuliert, dass jeweils ein anderer Aspekt der Befragung besonders betont wurde, um möglichst viele Personen anzusprechen. Der Beispieltext eines Reminders ist im Anhang in Abschnitt A.6 auf Seite 232 nachzulesen.

Nach Abschluss der Befragung wurde den Teilnehmern ihr gewähltes Incentive auf dem Postweg übersendet. Nach Fertigstellung der Auswertungen erhielten die Teilnehmer einen Best-Practice-Bericht sowie eine Checkliste mit Maßnahmen des verhaltensorientierten Controllings.

²²⁷ Vgl. für diese Vorgehensweise Dilts/Pence 2006b und Keil et alii 2003.

²²⁸ Vgl. beispielsweise Kalton 1983, S. 66.

4 Statistische Analysen und Diskussion der Ergebnisse

4.1 Datenaufbereitung und vorbereitende Analysen

In den folgenden Abschnitten werden die Aufbereitung der Daten sowie vorbereitende Analysen geschildert. Dazu zählt die Ermittlung der Rücklaufquote, die Prüfung der Normalverteilung und die Charakterisierung der Umfrageteilnehmer, Unternehmen und Projekte. Zusätzlich wurde bereits in dieser Phase eine Reihe von Tests zur Prüfung potentieller Biases durchgeführt. Diese Tests werden, gemeinsam mit den Fehlerprüfungen, die im Anschluss an die Entwicklung der Strukturmodelle vorgenommen wurden, in Abschnitt 4.6 auf Seite 182 vorgestellt.

4.1.1 Bestimmung und Diskussion der Rücklaufquote

Bevor die (Netto-)Rücklaufquote ermittelt werden kann, muss darüber entschieden werden, welche Datensätze nicht in die Analysen einbezogen werden. Grund für einen Ausschluss können beispielsweise fehlende oder fehlerhaft eingegebene Werte sein. Der Umgang mit fehlenden und fehlerhaften Daten wird im Folgenden erläutert.

4.1.1.1 Umgang mit fehlenden Daten

“Sooner or later (usually sooner), anyone who does statistical analysis runs into problems with missing data.”¹ Fehlende Daten stellen vor allem deshalb ein Problem dar, weil die Gefahr besteht, dass die Daten nicht zufällig fehlen, sondern die Nichtrespondenten in ihren Charakteristika von den Respondenten abweichen.² Der dadurch verursachte Fehler im Datensatz wird als Nonresponse-Bias bezeichnet.³ Die Höhe dieses Fehlers hängt von den tatsächlichen Unterschieden zwischen Respondenten und Nichtrespondenten in Bezug auf die untersuchten Fragen sowie von der Rücklaufquote ab.⁴ In Bezug auf fehlende Daten wird zwischen Unit-Nonresponse und Item-Nonresponse unterschieden. Im Fall der Unit-Nonresponse fehlt der gesamte Fragebogen, im Fall der Item-Nonresponse fehlen einzelne Werte.⁵

Unit-Nonresponse und Prüfung des Nonresponse-Bias

Als Unit-Nonresponse wird das vollständige Fehlen des Datensatzes einer Befragungseinheit (also in der Regel einer Person) bezeichnet.⁶ Die Bedrohung, dass sich die nichtantwortenden Personen in ihrer Einschätzung der untersuchten Thematik systematisch von den Antwortenden unterscheiden, ist insbesondere bei geringen Rücklaufquoten nicht zu unterschätzen.⁷ In der vorliegenden Studie liegt die Rücklaufquote bei 11,6 Prozent,⁸ so dass die Gefahr eines Nonresponse-Bias nicht ignoriert werden sollte.

Eine Methode, zu kontrollieren, ob systematische Unterschiede zwischen Respondenten und Nichtrespondenten vorliegen, besteht darin, die ersten und die letzten eingegangenen Antworten zu vergleichen. Dies basiert auf der Überlegung, dass die nicht Antwortenden den letzten

¹ Allison 2002, S. 1.

² So stellt Kalton fest: “In practice, however, it is dangerous to assume that the missing responses are missing at random; indeed, there are often good grounds for believing otherwise.”; Kalton 1983, S. 64. Ein offensichtlicher Unterschied besteht beispielsweise in dem Interesse für das Thema. Respondenten haben üblicherweise ein höheres Interesse an der Themenstellung als Nichtrespondenten; vgl. Armstrong/Overton 1977, S. 397. Diller schildert darüber hinaus die Erfahrung, dass Manager umso eher an einem Fragebogen teilnehmen, je bedürftiger sie in der entsprechenden Fragestellung sind. In der Thematik erfahrene und erfolgreiche Manager sind hingegen weniger auskunftsbereit, um ihren Informationsvorsprung nicht preiszugeben; vgl. Diller 2006, S. 616.

³ Vgl. beispielsweise Bosnjak/Tuten/Wittmann 2005, S. 490.

⁴ Vgl. beispielsweise Rogelberg/Luong 1998.

⁵ Vgl. Favre/Matei/Tillé 2005, S. 411.

⁶ Vgl. Bosnjak/Tuten/Wittmann 2005, S. 490.

⁷ Vgl. Kalton 1983, S. 66.

⁸ Vgl. Abschnitt 4.1.1.3 auf Seite 101.

Antwortenden ähnlicher sind als den früh Antwortenden.⁹ Für diesen Unit-Nonresponse-Bias-Test wurde jedes angeschriebene Sample in drei gleich große Gruppen (frühe Antworten, mittlere Antworten und späte Antworten) aufgeteilt. Anschließend wurde eine multivariate Analyse der Varianz (MANOVA) für alle Modellvariablen durchgeführt. Der Wilks- λ -Test erwies sich als nicht signifikant. Es liegen somit keine signifikanten Gruppenunterschiede vor, so dass der Nonresponse-Bias vernachlässigbar erscheint.

Zusätzlich wurde für alle Variablen die frühe Gruppe mit der späten Gruppe mittels eines beidseitigen t-Tests¹⁰ für unabhängige Stichproben verglichen.¹¹ Von den 76 metrisch skalierten Modellvariablen wiesen nur drei einen signifikanten Unterschied der Mittelwerte (auf einem Signifikanzniveau von 0,05) auf. Auch dies spricht dafür, dass kein nennenswerter Nonresponse-Bias vorliegt.

Zusätzlich wurden die in der Hoppenstedt-Datenbank verfügbaren Angaben genutzt, um mögliche Unterschiede zwischen den Antwortenden und den Nicht-Antwortenden aufzudecken. Wiederum wurden t-Tests durchgeführt und die folgenden Mittelwerte verglichen: Jahresumsatz, Bilanzsumme, Beschäftigtenanzahl und Jahresergebnis. Für keine der untersuchten Größen traten signifikante Unterschiede auf. Auch die Ergebnisse dieser Prüfung sprechen damit gegen das Vorliegen eines Unit-Nonresponse-Bias.¹²

Item-Nonresponse

Als Item-Nonresponse, Missing Items oder Missing Values werden einzelne fehlende Werte in Datensätzen bezeichnet. Diese können beispielsweise auf Grund der Wahl des „Weiß nicht“-Feldes oder durch Überspringen der Frage entstehen.¹³ In der Literatur werden zahlreiche Methoden vorgeschlagen und untersucht, um dem Problem des Item-Nonresponse zu begegnen.¹⁴ Dazu muss zunächst geklärt werden, um welche Art fehlender Daten es sich handelt, denn das Fehlen von Daten kann verschiedene Ursachen haben, die unterschiedliche Auswirkungen auf die statistische Analyse der Daten haben. Dabei geht es vor allem um die Frage, ob die Da-

⁹ Vgl. Armstrong/Overton 1977. Für Anwendungen dieses Tests vgl. beispielsweise Fisher/Katz 2000, S. 110, Griffis/Goldsby/Cooper 2003, S. 242, und Helasoja et alii 2002, S. 119 ff.

¹⁰ Für Grundlagen der Signifikanztests vgl. beispielsweise Mohr 1990.

¹¹ Für die genaue Vorgehensweise vgl. Pallant 2005, S. 255 ff.

¹² Dies umso mehr, als bei großen Samples üblicherweise auch kleine, unbedeutende Effekte statistisch signifikant werden können; vgl. Pallant 2005, S. 219.

¹³ Vgl. Schafer/Graham 2002, S. 148.

¹⁴ Vgl. beispielsweise Favre/Matei/Tillé 2005, Gilley/Leone 1991, Niyonsenga 1997.

ten zufällig fehlen (missing at random).¹⁵ In der Literatur hat sich die Klassifizierung in drei verschiedene Kategorien durchgesetzt.¹⁶

Missing completely at random (MCAR): Das Fehlen der Daten steht weder in einem Zusammenhang mit beobachteten noch mit fehlenden Werten.

Beispiel: X fehlt mit einer Wahrscheinlichkeit von 0,5.

Missing at random (MAR): Das Fehlen der Daten weist einen Zusammenhang mit beobachteten Werten auf, hängt jedoch *nicht* von den fehlenden Werten ab.

Beispiel: X fehlt, wenn $Y < 0$.

Not missing at random (NMAR): Das Fehlen der Daten hängt sowohl von den beobachteten als auch von den fehlenden Werten ab.

Beispiel: X fehlt, wenn $X < 0$.

Eine Standardmethode beim Umgang mit fehlenden Daten besteht im *Löschen des gesamten Datensatzes*¹⁷ des Teilnehmers, bei dem eine Antwort fehlt. Der Nachteil dieser Vorgehensweise besteht darin, dass vorhandene Daten wegen fehlender Daten gelöscht werden.¹⁸ Treten bei vielen Teilnehmern einzelne fehlende Werte auf, führt diese Methode zu einem erheblichen Datenverlust.¹⁹ Darüber hinaus kann das Löschen unvollständiger Datensätze zu verzerrten Schätzwerten führen, wenn die fehlenden Daten nicht MCAR sind.²⁰ Die vollständigen Datensätze sind dann nicht repräsentativ für die Population.²¹ In verschiedenen Simulationen wurde nachgewiesen, dass die listenweise Löschung zu mangelhaften Ergebnissen führt.²² Gemäß einer Studie von Bernaards/Sijtsma führt die listenweise Löschung teilweise sogar zu schlechteren Ergebnissen als ein zufälliges Ausfüllen der fehlenden Werte.²³

Eine andere Möglichkeit besteht darin, *die fehlenden Werte durch Schätzwerte zu ersetzen*, die auf der Basis der vorhandenen Daten ermittelt werden (data imputation).²⁴ Die folgenden sta-

¹⁵ Diese Diskussion geht zurück auf Rubin 1976.

¹⁶ Vgl. für die folgenden Definitionen und Beispiele Hippel 2004, S. 160 f.

¹⁷ Auch als Listwise Deletion, Case Deletion oder Complete-Case-Analysis bezeichnet.

¹⁸ Vgl. Arbuckle 2006, S. 273.

¹⁹ Vgl. Schafer/Graham 2002, S. 156.

²⁰ Vgl. Hippel 2004, S. 160.

²¹ Vgl. Schafer/Graham 2002, S. 157.

²² Vgl. beispielsweise Brown 1994 und Graham/Hofer/MacKinnon 1996.

²³ Vgl. Bernaards/Sijtsma 1999, S. 309.

²⁴ Vgl. Arbuckle 2006, S. 274.

tistischen Analysen werden dann mit dem vervollständigten Datensatz durchgeführt. Ein mögliches Imputations-Verfahren stellt beispielsweise der Expectation-Maximization-Algorithmus (EM)²⁵ von Dempster/Laird/Rubin dar,²⁶ der auf der Maximum-Likelihood-Methode (ML) basiert.²⁷ Gemäß einer Simulationsstudie von Bernaards/Sijtsma führt der EM-Algorithmus zu besseren Ergebnissen als alle fünf anderen untersuchten (Single-)Imputation-Varianten sowie die listenweise Löschung.²⁸ Entsprechen die fehlenden Daten der NMAR-Kategorie, so können auch ML-basierte Schätzverfahren zu verzerrten Ergebnissen führen. Allerdings weisen die ML-basierten Schätzwerte den geringsten Bias im Vergleich zu allen anderen Verfahren auf.²⁹ Somit sind ML-basierte Verfahren wie der EM-Algorithmus auf jeden Fall zu bevorzugen, wenn die Daten MAR oder NMAR entsprechen.

Für die vorliegende Studie ergab der MCAR-Test nach Little ein signifikantes Ergebnis. Dies bedeutet, dass die Daten nicht komplett zufällig fehlen.³⁰ Eine Imputation der fehlenden Daten mit dem EM-Algorithmus ist somit angebracht. Allerdings erscheint es fragwürdig, auch solche Datensätze zu imputieren, bei denen die Mehrzahl aller Items fehlt. Deshalb wurde ein zweistufiges Vorgehen gewählt. Für die Analysen wurden zunächst alle Fälle gelöscht (casewise deletion), bei denen mehr als die Hälfte der Items fehlt.³¹ Für die verbleibenden Datensätze wurde das EM-Schätzverfahren eingesetzt, um die fehlenden Werte zu berechnen. Für die meisten Variablen lag der zu schätzende Anteil bei unter fünf Prozent. Der höchste Anteil lag bei 23 Prozent.

²⁵ “Initially, random values are substituted for the missing data. In the *E* step, the missing values are updated given the factor scores, and the expected value of the covariance matrix given the factor loadings is calculated. In the *M* step the factor loadings and factor scores are updated based on the current estimate of the covariance matrix. These two steps are re-iterated until convergence of the likelihood occurs.”; Bernaards/Sijtsma 1999, S. 285; Hervorhebungen im Original.

²⁶ Vgl. Dempster/Laird/Rubin 1977.

²⁷ Vgl. Bernaards/Sijtsma 1999, S. 309.

²⁸ Vgl. Bernaards/Sijtsma 1999, S. 306. Hippel führt an, dass beispielsweise unter den in SPSS implementierten Verfahren zur Imputation von Missing Values das EM-Verfahren das einzige ist, das im MAR-Fall valide Resultate zeigt; vgl. Hippel 2004, S. 162. Allerdings wird auch die Implementierung dieses Verfahrens kritisiert, da SPSS keine Standardfehler ausweist und keine Residualvariationen berücksichtigt; vgl. Hippel 2004, S. 164.

²⁹ Vgl. beispielsweise Graham/Hofer/MacKinnon 1996, Little/Rubin 1989 und Muthén/Kaplan/Hollis 1987.

³⁰ Dies ist nicht ungewöhnlich für empirische Erhebungen. Muthén/Kaplan/Hollis stellen in diesem Zusammenhang fest, dass viele Forscher irrtümlich annehmen, ihre fehlenden Daten seien MCAR, obwohl sie dem MAR-Schema entsprechen; vgl. Muthén/Kaplan/Hollis 1987.

³¹ Ähnlich beispielsweise Ellingson/Sackett/Hough, die alle Datensätze eliminieren, bei denen mehr als ein Drittel der Items fehlt; vgl. Ellingson/Sackett/Hough 1999, S. 159.

4.1.1.2 Ausschluss von Random-Respondern

Eine potentielle Fehlerquelle stellen Personen dar, die den Fragebogen nicht ernsthaft beantworten. Dies kann beispielsweise der Fall sein, weil sich die Person nur den Fragebogen anschauen möchte oder an den Ergebnissen oder Incentives interessiert ist, ohne sich der Mühe des gewissenhaften Ausfüllens unterziehen zu wollen. Ein zufälliges Ausfüllen des Fragebogens wird auch als Random Responding bezeichnet. Eine Möglichkeit, dies zu kontrollieren, besteht in der Verwendung von Random Response Scales.³² Dabei werden Fragen eingesetzt, die jeder, der die Frage gelesen und verstanden hat, in der gleichen Weise beantworten muss. Ein Beispiel dafür ist die Aussage: „Ich wurde am 30. Februar geboren.“ Auch Ellingson/Sackett/Hough verwenden beispielsweise eine aus acht Items bestehende *Nonrandom Response Scale*, mit der zufällige Antworten identifiziert werden können.³³ Eine andere Möglichkeit besteht darin, die Antwortenden direkt zu fragen, ob sie den Fragebogen sorgfältig bearbeitet haben.³⁴

Ein Indiz für flüchtiges Ausfüllen ist die Bearbeitungszeit.³⁵ Darüber hinaus kann die Varianz der Antworten analysiert werden. Hat eine Person *stets* gleich geantwortet (also beispielsweise jede Skala ganz links angekreuzt), so ist die Wahrscheinlichkeit sehr gering, dass es sich um ernsthafte Antworten handelt. Im Vergleich zu den anderen Methoden haben die beiden letztgenannten Varianten den Vorteil, dass sie keine Items benötigen und damit den Fragebogen nicht verlängern. Darum werden für die vorliegende Studie diese Verfahren eingesetzt.

Die Antwortzeit wurde deshalb für jede Seite und pro Betrachter ermittelt und schließlich für jeden Teilnehmer die Ausfülldauer so normiert, dass die Werte zwischen null und eins liegen und der Durchschnitt bei 0,5. Anschließend wurden die Teilnehmer in der Reihenfolge ihrer genormten Ausfülldauer sortiert und die Verteilung betrachtet. Siebzehn Datensätze wurden schneller als in einem Viertel der durchschnittlichen Zeit ausgefüllt. Diese wurden hinsichtlich ihrer Antworten untersucht. Dabei wurde eine Person ermittelt, die alle Fragen ganz links angekreuzt hatte. Dieser Fall wurde eliminiert. Die anderen Fälle wiesen keine Auffälligkeiten auf. Zudem haben alle verbleibenden sechzehn Respondenten auch die freien Eingabefelder zum Fertigstellungsgrad des Projekts ausgefüllt. Einige der Personen hatten darüber hinaus Kommentare im Freitextfeld am Schluss des Fragebogens eingegeben. Diese Merkmale sprechen dafür, dass es auch in weniger als einem Viertel der Durchschnittszeit möglich war, den Fra-

³² Vgl. Beach 1989, S. 102.

³³ Vgl. Ellingson/Sackett/Hough 1999, S. 159.

³⁴ Vgl. Beach 1989, S. 102.

³⁵ Vgl. Archer 1997 und Archer/Elkins 1999, S. 408.

gebogen sorgfältig zu bearbeiten. Dementsprechend wurden keine Datensätze allein auf Grund der relativen Bearbeitungsdauer gelöscht.

Zur Kontrolle auf mangelnde Varianz der Antworten wurde für jeden Datensatz ermittelt, wie oft die acht verschiedenen Antwortmöglichkeiten (siebenstufige Skala plus „Weiß nicht“-Option) jeweils angekreuzt worden waren. Anschließend wurden die Datensätze der Reihe nach für jede der acht Möglichkeiten sortiert. Die Datensätze, bei denen die Antwortmöglichkeit sehr häufig oder gar nicht verwendet worden war, wurden genauer betrachtet, um mögliche Random-Responder zu identifizieren. Bei dieser Prozedur fielen insgesamt sieben Datensätze auf. Bei genauerer Untersuchung der einzelnen Antworten ließ sich jedoch nicht eindeutig bestimmen, ob es sich um Fälle von Random Responding handelt. Die Datensätze wurden deshalb im Sample belassen.

Neben den geschilderten Maßnahmen, um Random-Responder auszuschließen, wurde für die verbleibenden Datensätze untersucht, ob Aspekte des flüchtigen Ausfüllens zu einer systematischen Verzerrung der Antworten führen. Dies wurde mit Hilfe von Gruppenvergleichen geprüft und wird in Abschnitt 4.6.3 auf Seite 198 näher erläutert.

4.1.1.3 Berechnung der Rücklaufquote

Insgesamt wurden 10.056 Personen aus der Hoppenstedt-Datenbank angeschrieben. 1.908 E-Mails konnten nicht zugestellt werden. Von den verbleibenden 8.148 Personen füllten 949 Personen den Fragebogen aus. Die Rücklaufquote beträgt somit 11,6 Prozent. Zu Beginn des Fragebogens wurde erhoben, ob die Person Erfahrung mit abgebrochenen Projekten hat. 67 Personen gaben an, zwar mit neugestarteten Projekten (Relaunch) Erfahrung zu haben, jedoch nicht mit abgebrochenen. 205 Personen wiesen weder mit neugestarteten noch mit abgebrochenen Projekten Erfahrung auf. Diese Personen wurden von den folgenden Analysen ausgeschlossen. Somit verbleiben für die folgenden Analysen 677 Datensätze.

4.1.1.4 Beurteilung der Stichprobengröße

Der Stichprobenumfang sollte für kovarianzerklärende kausalanalytische Verfahren nach Möglichkeit mindestens 200 bis 250 Fälle betragen.³⁶ Eine grobe Abschätzung erlaubt auch die Empfehlung von Bentler/Chou, wonach mindestens fünfmal so viele Fälle bzw. Datensätze vorliegen sollten, wie die Anzahl der freien (d. h. der zu schätzenden) Modellparameter beträgt.³⁷

³⁶ Vgl. Homburg/Klarmann 2006, S. 733.

³⁷ Vgl. Bentler/Chou 1987, S. 91.

Das umfangreichste Modell, das in der vorliegenden Arbeit geschätzt wird,³⁸ weist 78 freie Parameter auf. Die Anzahl der Datensätze von 677 erfüllt somit die Forderungen der Literatur und lässt auch Gruppenvergleiche zu, bei denen das Sample geteilt wird.³⁹

Ein Indikator für die Ex-post-Beurteilung der Stichprobengröße für ein gegebenes Modell (nach der Datenerhebung und Modellschätzung) stellt der RMSEA dar.⁴⁰ Für die untersuchten Modelle (vgl. Abschnitt 4.4.1 auf Seite 148) ist der RMSEA zufriedenstellend. Eine weitere Möglichkeit, um zu prüfen, ob der Stichprobenumfang ausreichend groß ist, besteht im Durchführen von Stabilitätschecks.⁴¹ Die Stabilitätschecks ergeben für den vorliegenden Datensatz bei zufälligem Entfernen von zehn Prozent der Datensätze keine nennenswerten Veränderungen in den Ergebnissen.

Insgesamt ist also festzuhalten, dass der Umfang der Stichprobe zuverlässige Ergebnisse der statistischen Analysen erwarten lässt.

4.1.1.5 Suche nach fehlerhaften Werten

Ausreißer können erhebliche Auswirkungen auf die statistischen Analysen verursachen und im Rahmen von SEM insbesondere Gütemaße, Schätzwert und Standardfehler verfälschen.⁴² Entsprechend den Empfehlungen von Pallant wurden die Daten aller Variablen auf Ausreißer und sonstige mögliche Eingabefehler untersucht.⁴³

Dabei wurde für jede Variable einzeln geprüft, ob der kleinste und der größte Wert plausibel sind und ob die Anzahl der fehlenden Werte im erwarteten Rahmen liegt. Für metrisch skalierte Daten wurden zusätzlich Mittelwert und Standardabweichung auf Plausibilität überprüft. Bei der Suche nach Ausreißern wurde beispielsweise entdeckt, dass eine Angabe zum „Grad der Fertigstellung im Vergleich zur Planung“ bei 505 Prozent, d. h. über 100 Prozent, lag. Es besteht die Möglichkeit, dass damit gemeint war, dass bis zum Abbruch fünfmal so viel fertiggestellt wurde. Wahrscheinlicher jedoch erscheint es, dass die Angabe ein Versehen darstellt.

³⁸ Vgl. Abbildung 4.21 auf Seite 194.

³⁹ Allerdings sollte bei 677 Datensätzen die damit ermittelbare Genauigkeit der statistischen Analysen nicht überschätzt werden: “Most researchers report correlations rounded to two digits. If this is the required accuracy for knowledge of a correlation, we can compute the sample size needed so that the sample correlation will come out correctly to two digits. The sample size is [...] $N = 153.669$.”; Hunter 2001, S. 149 f.

⁴⁰ Vgl. Bentler/Chou 1987. Zur Erläuterung des RMSEA vgl. Seite 119.

⁴¹ Vgl. Abschnitt 4.6.4 auf Seite 200.

⁴² Vgl. West/Finch/Curran 1995, S. 61.

⁴³ Vgl. Pallant 2005, S. 40 ff.

Dafür sprechen vor allem die folgenden Gründe: Kein anderer Fragebogen enthielt eine Angabe von über 100 Prozent. Die Angaben des entsprechenden Datensatzes für den Prozentsatz für Manntage, Kosten und Kalenderzeit im Vergleich zur Planung lauten 65, 75 und 70 Prozent. Es erscheint unwahrscheinlich, dass ein Projekt mit lediglich zwei Dritteln der veranschlagten Ressourcen fünfmal so viel Ergebnisse erzielt wie ursprünglich geplant, dann jedoch abgebrochen wird. Die Angabe wurde deshalb als fehlender Wert definiert.

4.1.2 Prüfung auf Normalverteilung

Die klassischen Methoden für die Analyse von Strukturgleichungsmodellen basieren auf der Annahme multivariater Normalverteilung.⁴⁴ Empirische Daten in den Sozial- und Verhaltenswissenschaften weisen jedoch typischerweise starke Abweichungen von der Normalverteilung auf.⁴⁵ Backhaus/Blechsmidt/Eisenbeiß empfehlen deshalb, vor der Modellschätzung die Verteilungseigenschaften der betrachteten Daten zu analysieren.⁴⁶ Gemäß einer Studie von Breckler, der 72 Artikel analysiert, in denen die SEM-Methode verwendet wird, erwähnen jedoch nur 19 Prozent der Autoren die Normalverteilungsannahme und nur 10 Prozent testen explizit eine mögliche Verletzung dieser Grundannahme.⁴⁷

Für den vorliegenden Datensatz wurde zunächst die univariate Normalverteilung mit Hilfe des Kolgomorov-Smirnov-Tests geprüft. Der Kolgomorov-Smirnov-Test auf Normalverteilung mit Signifikanzkorrektur nach Lilliefors ergibt für alle im Modell verwendeten Variablen des Datensatzes einen Signifikanzwert von unter 0,05. Die Annahme von normalverteilten Daten wird damit abgelehnt. Dies ist allerdings kein überraschendes Ergebnis, sondern für große Samples der Regelfall,⁴⁸ da die überwiegende Mehrheit von Erhebungsdaten nicht das Kriterium der multivariaten Normalverteilung erfüllt.⁴⁹ Da Tests auf Schiefe und Kurtosis bei großen Samples in der Regel zu sensitiv sind, wird in der Literatur empfohlen, die Verteilung anhand eines Diagramms (beispielsweise eines Histogramms) zu beurteilen.⁵⁰ Deshalb wurden für alle Variablen je ein Histogramm, ein Q-Q-Diagramm und ein trendbereinigtes Q-Q-Diagramm

⁴⁴ Vgl. Yuan/Bentler/Chan 2004, S. 421.

⁴⁵ Vgl. beispielsweise Pallant 2005, S. 52 und 58, und Yuan/Bentler/Chan 2004, S. 421.

⁴⁶ Vgl. Backhaus/Blechsmidt/Eisenbeiß 2006, S. 715.

⁴⁷ Vgl. Breckler 1990.

⁴⁸ Vgl. Pallant 2005, S. 57.

⁴⁹ Vgl. Byrne 2001, S. 268.

⁵⁰ Vgl. beispielsweise Tabachnick/Fidell 2001, S. 73.

erstellt und betrachtet.⁵¹ Die Diagramme zeigen für einen Teil der Variablen eine Verteilung, die in etwa der Normalverteilung entspricht. Andere Variablen weichen sehr deutlich von der Normalverteilung ab. Homburg/Klarmann schätzen ein, dass keine wesentlichen Auswirkungen auf die Schätzparameter zu befürchten sind, wenn die Schiefe und Wölbung der meisten Variablen zwischen +1 und -1 liegt.⁵² Für die vorliegende Arbeit bewegen sich die Schiefemaße für alle Items mit siebenstufigen Antwortskalen in dem genannten Bereich. Der *Betrag* der meisten Kurtosis-Maße liegt unter eins, einzelne Wert überschreiten eins, wobei der Maximalwert 1,4 beträgt. Für die Felder mit Prozentangaben (bei dem Konstrukt zum Anteil der investierten Ressourcen, vgl. Absatz 4.3.1.3 auf Seite 126) liegt die Schiefe zwischen -0,3 und +3,1. Die Kurtosis liegt zwischen -0,9 und +20,3. Da dies jedoch nur vier Variablen betrifft, ist gemäß Homburg/Klarmann für das Gesamtmodell keine wesentliche Verzerrung zu erwarten.⁵³

Eine univariate Normalverteilung der *einzelnen Indikatorvariablen* ist eine notwendige, jedoch keine hinreichende Bedingung für das Vorliegen einer *multivariaten* Normalverteilung.⁵⁴ Der Mardia-Koeffizient der multivariaten Kurtosis stellt eine Möglichkeit dar, die multivariate Normalverteilung zu prüfen. Weicht der Wert signifikant von null ab, bedeutet dies, dass eine signifikante Wölbung vorliegt, die Daten also von der multivariaten Normalverteilung abweichen.⁵⁵

Für den vorliegenden Datensatz wurde der Mardia-Koeffizient der multivariaten Kurtosis für das Gesamtmodell mit dem Index der Maßnahmen (vgl. Abbildung 4.12 auf Seite 177) ermittelt. Der Wert liegt bei 164 und ist deutlich signifikant (t-Wert = 60,6). Die Bedingung der multivariaten Normalverteilung ist demnach *nicht* erfüllt. Dies ist jedoch kein Grund, auf die Analyse mit Hilfe von Strukturgleichungsmodellen zu verzichten. „Zahlreiche Simulationsstudien haben allerdings gezeigt, dass Abweichungen von der multivariaten Normalverteilung höchstens zu einer geringen Verzerrung der geschätzten Parameter führen.“⁵⁶

Weichen die Daten von der multivariaten Normalverteilung ab, kann dies vor allem zwei Folgen haben: Der χ^2 -Test lehnt zu viele Modelle ab und die Tests der Schätzwerte ergeben zu

⁵¹ Vgl. hierzu beispielsweise Pallant 2005, S. 53 ff.

⁵² Vgl. Homburg/Klarmann 2006, S. 736. West/Finch/Curran nennen als Beispielgrößen für substantiell nicht normalverteilte Daten eine Schiefe von 2 und eine Wölbung von 7; vgl. West/Finch/Curran 1995, S. 74.

⁵³ Homburg/Klarmann 2006, S. 736.

⁵⁴ Vgl. Backhaus et alii 2006, S. 715.

⁵⁵ Vgl. Backhaus et alii 2006, S. 715, und West/Finch/Curran 1995, S. 61.

⁵⁶ Homburg/Klarmann 2006, S. 736. Für eine kurze Zusammenfassung der Ergebnisse von Studien über die Wirkung nicht normalverteilter Daten auf ML- und GLS-Schätzung vgl. West/Finch/Curran 1995, S. 62 ff.

viele signifikante Ergebnisse.⁵⁷ Deshalb sollte in diesem Fall geprüft werden, wie robust die Ergebnisse der Analysen sind.⁵⁸

Eine Möglichkeit, die Robustheit der Ergebnisse im Fall nicht-normalverteilter Daten zu prüfen, bietet die Bootstrap-Methode. Das Bootstrap-Verfahren kann dabei, ebenso wie das Standard-Verfahren, auf unterschiedlichen Schätzverfahren basieren, beispielsweise dem Maximum-Likelihood-Verfahren (ML) oder dem Generalized-Least-Squares-Verfahren (GLS). Die folgenden Ausführungen beziehen sich immer auf das ML-Verfahren, wenn nicht ausdrücklich ein anderes benannt wird. Das Bootstrap-Verfahren geht auf Efron zurück und stellt ein Resampling-Verfahren dar.⁵⁹ Das Originalsample wird dabei als repräsentativ für die Grundgesamtheit betrachtet. Beim Bootstrap-Verfahren werden mehrere (beispielsweise 500) neue Samples mit der gleichen Größe wie das Originalsample generiert, indem aus dem Originalsample zufällig Datensätze *mit Zurücklegen* gezogen werden.⁶⁰ Wenn die Daten normalverteilt sind, unterliegen die Standardfehler beim Bootstrap-Verfahren einer größeren Verzerrung als beim Standard-Verfahren. Wenn die Verteilung jedoch nicht normal ist, sind die Bootstrap-Schätzungen weniger verzerrt als die Standard-Schätzungen.⁶¹

Wendet man mit AMOS das Bootstrap-Verfahren an, erhält man *ein* Set von Schätzungen, jedoch *zwei* Sets von Standardfehlern. Das erste Set entspricht den üblicherweise erstellten Schätzungen und basiert auf der Annahme der Normalverteilung. Das zweite Set wird über das Bootstrap-Verfahren ermittelt, ohne dabei auf die Annahme der Normalverteilung zurückzugreifen.⁶² Ein Vergleich der beiden Sets von Standardfehlern gibt Aufschluss darüber, wie stark sich die Abweichung von der Normalverteilung auf die Zuverlässigkeit der Schätzungen auswirkt.

Zur Prüfung der Auswirkungen der Abweichung von der multivariaten Normalverteilung wurde das Bootstrap-Verfahren mit 500 Samples angewendet.⁶³ Dabei wurde das in Abbildung 4.12 auf Seite 177 dargestellte Modell verwendet. Tabelle 4.1 stellt die Unterschiede der Standardfehler der Pfadkoeffizienten zwischen Standard-ML-Verfahren und Bootstrapping-Verfahren

⁵⁷ Vgl. West/Finch/Curran 1995, S. 59.

⁵⁸ Vgl. West/Finch/Curran 1995.

⁵⁹ Vgl. Efron 1979.

⁶⁰ Vgl. Byrne 2001, S. 268.

⁶¹ Vgl. Byrne 2001, S. 270, und Ichikawa/Konishi 1995.

⁶² Vgl. Byrne 2001, S. 270.

⁶³ Vgl. Byrne 2001, S. 281.

Tabelle 4.1: Standardfehler der Pfadkoeffizienten beim Bootstrap-Verfahren

unabhängige Variable		abhängige Variable	R	t	ML-SE	BS-SE	Δ
Wahrnehmungsschwelle	→	Projekteskalation	0,17	2,79	0,061	0,070	15 %
selektive Wahrnehmung	→	Projekteskalation	-0,05	-1,00	0,050	0,063	26 %
Self-Justification-Druck	→	Projekteskalation	-0,02	0,60	0,040	0,044	10 %
Sunk-Cost-Effekt	→	Projekteskalation	0,27	5,61	0,049	0,054	10 %
übermäßiger Optimismus	→	Projekteskalation	0,19	4,72	0,039	0,044	13 %
Index aller Maßnahmen	→	Wahrnehmungsschwelle	0,00	0,12	0,002	0,002	0 %
Index aller Maßnahmen	→	selektive Wahrnehmung	-0,01	-3,85	0,002	0,003	50 %
Index aller Maßnahmen	→	Self-Justification-Druck	0,00	-0,77	0,003	0,003	0 %
Index aller Maßnahmen	→	Sunk-Cost-Effekt	0,00	0,49	0,003	0,003	0 %
Index aller Maßnahmen	→	übermäßiger Optimismus	0,00	0,23	0,003	0,003	0 %
Index aller Maßnahmen	→	Projekteskalation	-0,01	-2,62	0,002	0,002	0 %

R = nicht standardisiertes Regressionsgewicht, t = t-Wert

ML-SE = Maximum Likelihood Standard Error

BS-SE = Bootstrap Standard Error, Δ = prozentuale Abweichung

dar. Einige der im Bootstrap-Verfahren ermittelten Standardfehler liegen etwas höher als die Standardfehler gemäß Standard-ML-Verfahren.

Im nächsten Schritt wird geprüft, welche Effekte zwischen den Variablen bei der Anwendung des Bootstrap-Verfahrens signifikant sind. Tabelle 4.2 zeigt die untere und die obere Grenze der Regressionsgewichte für ein 90-Prozent-Konfidenzintervall an. Mit 90 Prozent Wahrscheinlichkeit liegen die Regressionsgewichte also zwischen diesen Grenzen.

Tabelle 4.2: Konfidenzintervalle der Pfadkoeffizienten beim Bootstrap-Verfahren

unabhängige Variable		abhängige Variable	R	Lo	Hi	p
Wahrnehmungsschwelle	→	Projekteskalation	0,17	0,051	0,282	0,011
selektive Wahrnehmung	→	Projekteskalation	-0,05	-0,147	0,051	0,447
Self-Justification-Druck	→	Projekteskalation	-0,02	-0,055	0,093	0,642
Sunk-Cost-Effekt	→	Projekteskalation	0,27	0,194	0,371	0,003
übermäßiger Optimismus	→	Projekteskalation	0,19	0,119	0,269	0,003
Index aller Maßnahmen	→	Wahrnehmungsschwelle	0,00	-0,007	0,001	0,162
Index aller Maßnahmen	→	selektive Wahrnehmung	-0,01	-0,014	-0,005	0,004
Index aller Maßnahmen	→	Self-Justification-Druck	0,00	-0,008	0,003	0,445
Index aller Maßnahmen	→	Sunk-Cost-Effekt	0,00	-0,003	0,006	0,536
Index aller Maßnahmen	→	übermäßiger Optimismus	0,00	-0,004	0,006	0,830
Index aller Maßnahmen	→	Projekteskalation	-0,01	-0,010	-0,002	0,021

R = nicht standardisiertes Regressionsgewicht

Lo = untere Grenze des 90-%-Konfidenzintervalls

Hi = obere Grenze des 90-%-Konfidenzintervalls

p = Signifikanzniveau, bei dem Lo = 0 wäre

Umschließen die untere und obere Grenze in Tabelle 4.2 *nicht* den Nullpunkt, so wird die Hypothese einer signifikanten Wirkung zwischen den Faktoren gestützt. Umfassen die Grenzen hin-

gegen den Nullpunkt, wie dies bei den Wirkungen selektive Wahrnehmung → Projekteskalation, Self-Justification-Druck → Projekteskalation sowie aller Wirkungen des Maßnahmenindex (mit Ausnahme von Index aller Maßnahmen → selektive Wahrnehmung) der Fall ist, kann die Wirkung *nicht* auf dem 90-Prozent-Konfidenzintervall bestätigt werden.⁶⁴

Das Bootstrap-Verfahren kommt somit zum gleichen Ergebnis wie das Standard-ML-Verfahren:⁶⁵ Die Projekteskalation wird signifikant durch die Wahrnehmungsschwelle, den Sunk-Cost-Effekt, den übermäßigen Optimismus und den Index der Summe aller Maßnahmen beeinflusst. Die Summe der Maßnahmen wirkt sich darüber hinaus auf die selektive Wahrnehmung aus. Alle anderen Effekte sind für das Gesamtsample nicht signifikant. Da beide Verfahren zu den gleichen Ergebnissen gelangen, können die Abweichungen von der Normalverteilung für die geschätzten Modellparameter vernachlässigt werden.

4.1.3 Charakterisierung der Respondenten, Unternehmen und Projekte

Die Untersuchung richtete sich vor allem an mittlere und große Unternehmen und war branchenübergreifend. Die beteiligten Branchen gehen aus Abbildung 4.1 hervor, wobei die Bezeichnungen auf dem Branchencode des europäischen Standards von Eurostat: Statistisches Amt der Europäischen Gemeinschaften basieren.⁶⁶

Der Großteil der Projekte bestand aus Optimierungen/Prozessverbesserungen, Entwicklungsprojekten und Implementierungsprojekten. Diese drei Projektarten machen gemeinsam knapp zwei Drittel aller betrachteten Projekte aus. In Bezug auf die Investitionssumme ist festzustellen, dass im betrachteten Sample einige sehr große Projekte sind, die deutlich über dem Durchschnitt liegen. Deshalb werden in Abbildung 4.1 nicht nur die arithmetischen Mittelwerte angegeben, die stärker durch Ausreißer verzerrt werden, sondern auch die Mediane.⁶⁷

⁶⁴ Vgl. Byrne 2001, S. 283.

⁶⁵ Vgl. Abbildung 4.12 auf Seite 177.

⁶⁶ Vgl. Eurostat: Statistisches Amt der Europäischen Gemeinschaften 2007. Zur Verbesserung der Übersichtlichkeit wurden einige Branchen zu Oberklassen zusammengefasst.

⁶⁷ Für eine Analyse, inwiefern sich große Projekte hinsichtlich der Wirkungen von kognitiven Faktoren und Gegenmaßnahmen auf die Projekteskalation von den übrigen Projekten unterscheiden, vgl. Abschnitt 4.6.3 auf Seite 198.

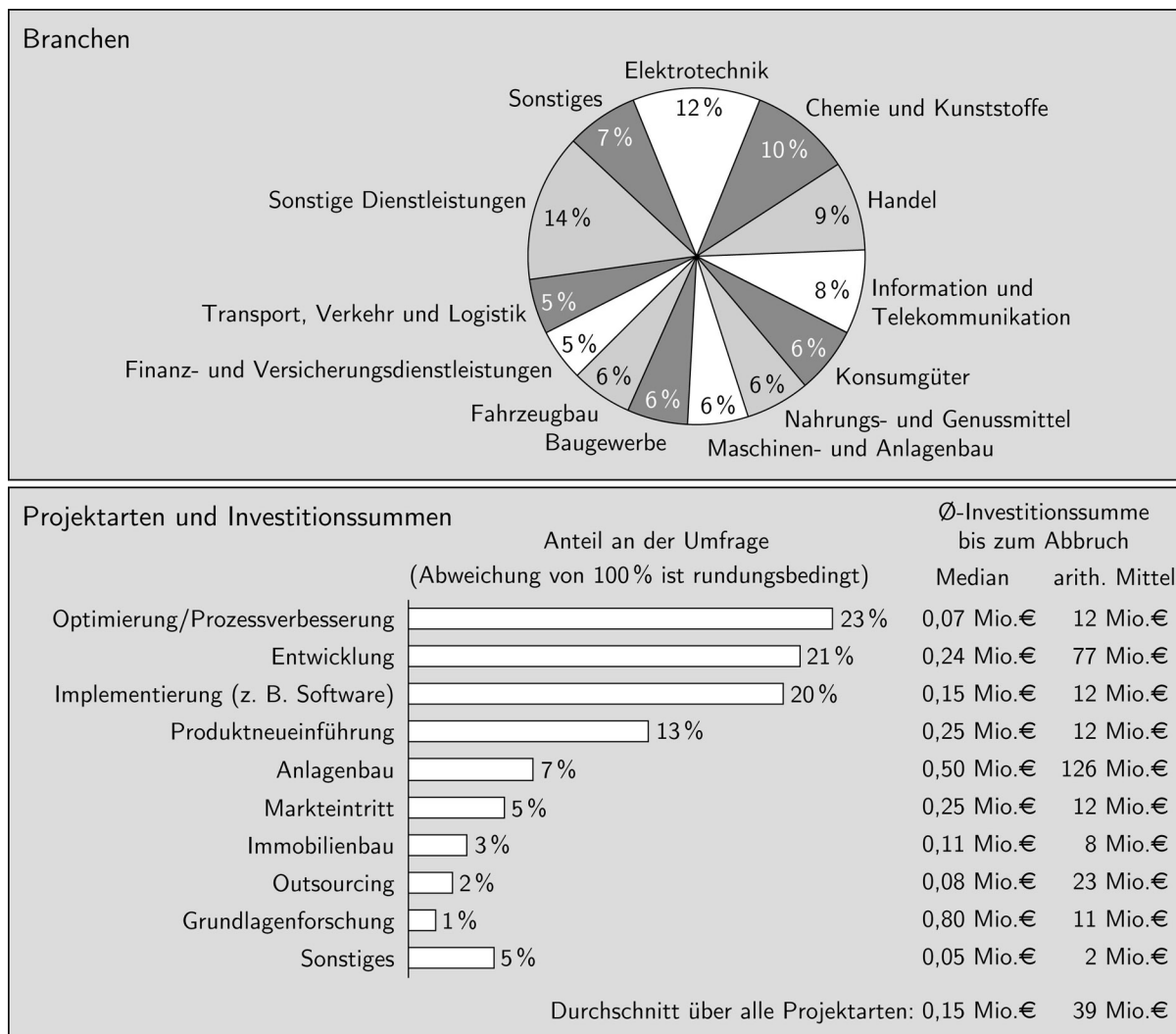


Abbildung 4.1: Charakteristika der Projekte

Der arithmetische Mittelwert der investierten Kosten liegt bei 39 Mio. Euro. Der arithmetische Mittelwert der *geplanten* Projektkosten beträgt 68 Mio. Euro (bei einem geplanten Einsatz von etwa 1.100 Manntagen und einer geplanten Dauer von zwei Jahren). Bis zum Abbruch wurden im Durchschnitt etwa zwei Drittel der ursprünglich geplanten Gesamtsumme investiert.⁶⁸

Abbildung 4.2 stellt die Positionen der Umfrageteilnehmer und ihre Rolle im Projekt dar. Von den Teilnehmern arbeiten 42 Prozent im Bereich Controlling/Finanzen und ein Drittel in technischen Führungspositionen (insbesondere Forschung und Entwicklung). Die restlichen Befragten üben Tätigkeiten im General Management oder als Berater aus.

⁶⁸ Für eine Analyse, inwiefern sich die Top zehn Prozent der Projekte hinsichtlich des Ressourceneinsatzes von den übrigen Projekten unterscheiden, vgl. Abschnitt 4.5.3 auf Seite 178.

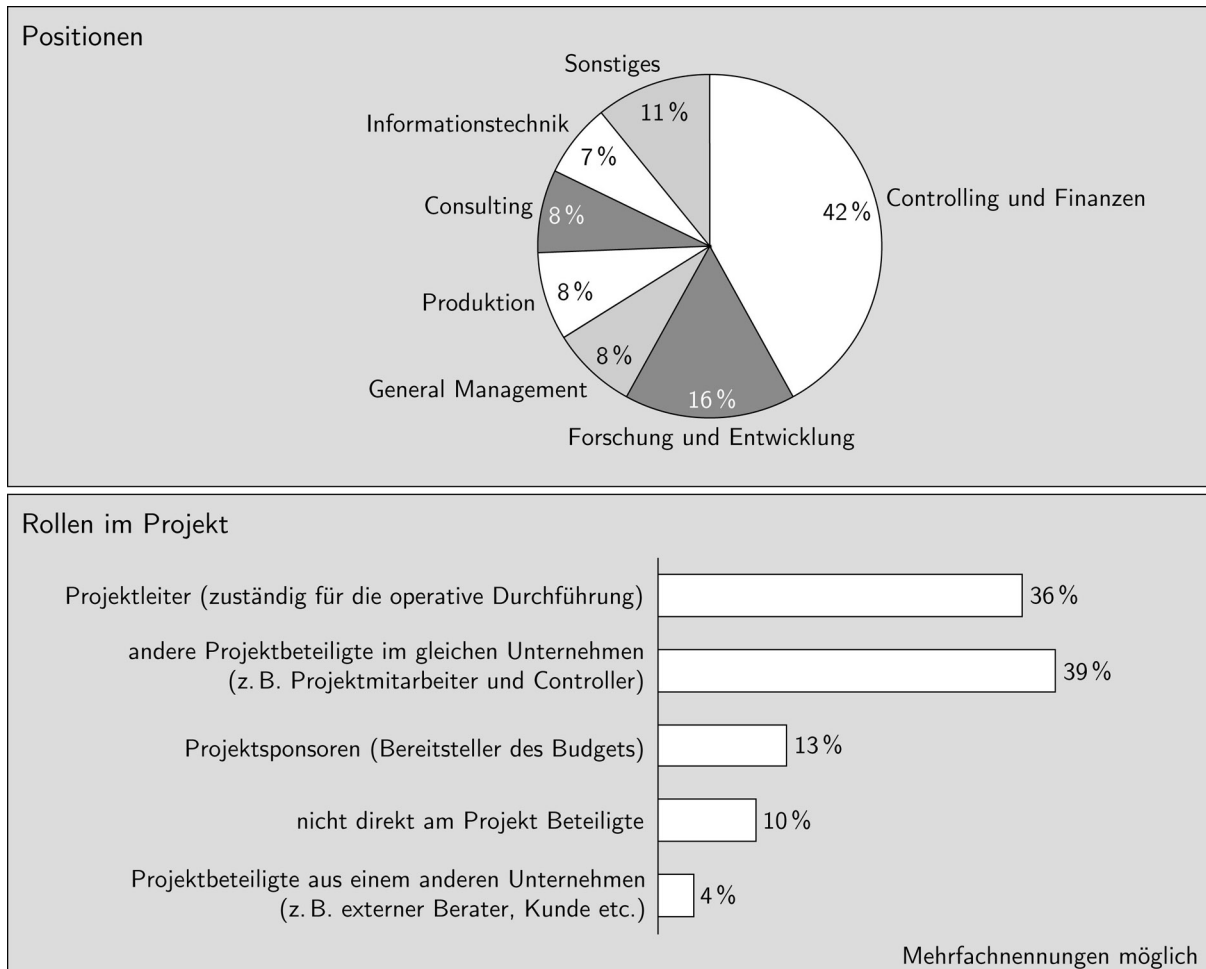


Abbildung 4.2: Charakteristika der Umfrageteilnehmer

Von den Befragten war etwas mehr als ein Drittel als Projektleiter auch für die operative Steuerung des abgebrochenen Projekts tätig. 39 Prozent waren als Projektmitarbeiter oder Controller am Projekt beteiligt. 13 Prozent der Befragten hatten die Rolle des Projektsponsors inne, waren also für die Bereitstellung des Budgets verantwortlich. Die Rollen der verbleibenden Umfrageteilnehmer setzen sich aus Beratern, Kunden sowie nicht aktiv am Projekt beteiligten Beobachtern zusammen.

Die Befragten verfügten zum Zeitpunkt des Projektabbruchs durchschnittlich über mehr als vierzehn Jahre Berufserfahrung und waren mehr als acht Jahre im betreffenden Unternehmen beschäftigt. Etwa 40 Prozent der Befragten waren für die Entscheidung über Abbruch oder Fortsetzung des Projekts verantwortlich.

4.2 Methodische Hintergründe zur kovarianzbasierten Analyse von Strukturgleichungsmodellen

4.2.1 Auswahl des statistischen Analyseverfahrens

Statistische Analyseverfahren tragen nur dann zum Erkenntnisgewinn bei, wenn sie auf eine angemessene Problemstellung angewendet werden – dies ist oftmals auch bei anspruchsvollen Verfahren *nicht* der Fall. So konstatiert beispielsweise Diller: „Das süße Gift der vermeintlich intellektuellen Spitzenleistung einer Kausalmodell-Überprüfung kann bei solchen Fehlanwendungen zu viele und zu schwerwiegende langfristige Auswirkungen auf die Fortentwicklung unserer Theorien, aber auch auf das Ansehen des Forschers und der gesamten Scientific Community zeitigen.“⁶⁹ Die Auswahl des Verfahrens soll deshalb im Folgenden begründet werden.

Die Forschungsfragen⁷⁰ der vorliegenden Arbeit zielen auf eine Erklärung von Abhängigkeitsstrukturen ab. Das am *häufigsten* angewendete Verfahren zur Analyse von Abhängigkeitsstrukturen stellt die Regressionsanalyse dar.⁷¹ Die multiple Regressionsanalyse wird jedoch in der vorliegenden Arbeit *nicht* eingesetzt, da sie zwei wesentliche Schwächen aufweist:⁷²

(1) Die Regressionsanalyse kann Messfehler nicht berücksichtigen, was die Ergebnisse und Interpretationen stark verfälschen kann. (2) Die Regressionsanalyse ist meist nicht in der Lage, komplexe Kausalstrukturen wie gestufte Abhängigkeiten (kausale Ketten) und wechselseitige Abhängigkeiten zu modellieren.

Die genannten Schwächen der Regressionsanalyse treten bei der Kovarianzstrukturanalyse (auch als Kausalanalyse bezeichnet) nicht auf.⁷³ Unter den multivariaten Verfahren, die sich mit Abhängigkeitsstrukturen befassen, gilt die Kovarianzstrukturanalyse deshalb gegenwärtig als das leistungsfähigste Verfahren.⁷⁴ Varianzbasierte Schätzungen von Strukturgleichungsmodellen mit dem Partial-Least-Square-Verfahren (PLS) werden in der Literatur als weniger

⁶⁹ Diller 2006, S. 616.

⁷⁰ Vgl. Abschnitt 1.3 auf Seite 7.

⁷¹ Vgl. Homburg 1992b, S. 499.

⁷² Vgl. für die folgenden Aussagen Homburg 1992b, S. 499.

⁷³ Für die Grundlagen der Kovarianzanalyse vgl. beispielsweise Wildt/Ahtola 1978.

⁷⁴ Vgl. Diller 2006, S. 611, Homburg 1992b, S. 499.

leistungsfähig beurteilt. So kommen beispielsweise Scholderer/Balderjahn zu dem Schluss: „PLS sollte nach unserer Auffassung insofern nur eingesetzt werden, wenn ein Einsatz von LISREL [d. h. einem kovarianzbasierten Verfahren] definitiv nicht möglich ist.“⁷⁵ Dies ist dann der Fall, wenn die Stichprobe klein ist und das Strukturmodell sehr viele Variablen beinhaltet sowie komplexe Strukturen aufweist.⁷⁶ Balderjahn beziffert den relevanten Stichprobenumfang mit 200 und nennt als Beispiel für einen gerechtfertigten PLS-Einsatz ein Modell mit 68 Variablen.⁷⁷ In der vorliegenden Arbeit liegt der Stichprobenumfang deutlich höher als 200 und die Anzahl der Modellvariablen deutlich unter 68, so dass der Einsatz von PLS nicht notwendig und gemäß der Argumentation von Scholderer/Balderjahn auch nicht empfehlenswert ist.

Die Leistungsfähigkeit der Kovarianzanalyse spiegelt sich auch in ihrer zunehmenden Nutzung wider.⁷⁸ Insbesondere für die Konzeptionalisierung und Operationalisierung *latenter* Konstrukte ist die Kovarianzstrukturanalyse besonders geeignet.⁷⁹ Im Rahmen von Kausalanalysen wird zwischen direkt messbaren Variablen (wie beispielsweise einem Verkaufspreis) und nicht direkt messbaren Variablen (wie beispielsweise der Motivation eines Mitarbeiters) unterschieden.⁸⁰ Letztere werden als latente Variablen bzw. hypothetische Konstrukte bezeichnet.⁸¹ Werden Fragestellungen untersucht, die nur mit latenten Variablen geprüft werden können, bietet sich die Nutzung von *Strukturgleichungsmodellen* an.⁸² Dies ist bei der vorliegenden Arbeit der Fall.⁸³

Aus der kurzen Diskussion ergibt sich die Schlussfolgerung, dass für die vorliegende Studie die Kovarianzstrukturanalyse das aus wissenschaftlicher Sicht am besten geeignete Analyseverfahren darstellt. Deshalb wird in dieser Arbeit die Kovarianzstrukturanalyse verwendet. Da auch Kovarianzanalysen nicht die Wirkungsrichtung eines Effekts zwischen zwei Variablen ermit-

⁷⁵ Scholderer/Balderjahn 2006, S. 67.

⁷⁶ Vgl. Albers/Hildebrandt 2006, S. 15 ff. und Balderjahn 2008, S. 267.

⁷⁷ Vgl. Balderjahn 2008, S. 267 ff.

⁷⁸ Vgl. Homburg/Baumgartner 1995. Für Rechenbeispiele zur Kovarianzstrukturanalyse vgl. Homburg 1992a, S. 541 ff.

⁷⁹ So erläutern beispielsweise Homburg/Klarmann: „Als erste und bislang einzige Methode erlaubt die Kovarianzstrukturanalyse die Analyse von komplexen Abhängigkeitsstrukturen bei gleichzeitiger Berücksichtigung der Problematik, dass viele theoretisch interessante Phänomene nicht direkt messbar sind.“; Homburg/Klarmann 2006, S. 727.

⁸⁰ Für eine ausführliche Erklärung vgl. Backhaus et alii 2006, S. 339 f.

⁸¹ Vgl. Backhaus et alii 2006, S. 339.

⁸² Vgl. Backhaus et alii 2006, S. 338.

⁸³ Vgl. Abschnitt 3.2.2.5 auf Seite 59. Die Operationalisierung wurde in Kapitel 3.2.2 auf Seite 52 erläutert. In Kapitel 4.3 ist der genaue Wortlaut für jedes Item nachzulesen.

teln können, ist es von entscheidender Bedeutung, die Kausalbeziehungen zwischen den Variablen *vorab* zu postulieren.⁸⁴ Eine häufig genannte Forderung besteht deshalb darin, *zuerst theoretisch* fundierte Hypothesen zu bilden und *erst danach konfirmatorisch* diese Hypothesen zu prüfen.⁸⁵ Dieser Forderungen wird in der vorliegenden Arbeit gefolgt. Die Hypothesen wurden in Kapitel 2.2 auf Seite 21 ff. theoretisch fundiert.

4.2.2 Kriterien für die Gütebeurteilung der Messmodelle

Vor drei Jahrzehnten kritisierte Jacoby: “Most of our measures are only measures, because someone says they are, not because they have been shown to satisfy standard measurement criteria.”⁸⁶ Dies hat sich inzwischen gewandelt. „Mittlerweile gehört es zum *state of the art*, latente Variablen einem Skalenbeurteilung und -bereinigungsprozess zu unterwerfen.“⁸⁷ Bevor dies in Abschnitt 4.3 auf Seite 121 ff. erfolgt, sollen die Kriterien kurz dargestellt werden, die der Skalenbeurteilung zu Grunde liegen.

4.2.2.1 Kriterien zur Prüfung der Reliabilität

Der Begriff Reliabilität bezeichnet die Genauigkeit einer Messung. Ein Testverfahren ist umso reliabler, je geringer die zufälligen Messfehler der erhaltenen Testwerte sind.⁸⁸ Die Reliabilität kann sowohl (1) auf der Ebene einzelner Indikatoren betrachtet werden als auch (2) auf der Ebene der Faktoren.

(1) Die *Indikatorreliabilität* entspricht dem Quadrat der (standardisierten) Faktorladung des jeweiligen Indikators.⁸⁹ Dieser Wert kann zwischen null und eins liegen, wobei die Reliabilität bei einem Wert nahe eins sehr hoch und bei einem Wert nahe null sehr niedrig ist. In Bezug

⁸⁴ Vgl. Backhaus et alii 2006, S. 348.

⁸⁵ Vgl. für diese Forderung beispielsweise Backhaus et alii 2006, S. 338, Birnberg 1973, S. 75 f., und Birnberg/Nath 1967, S. 474.

⁸⁶ Jacoby 1978, S. 91.

⁸⁷ Fassott 2006, S. 69; Hervorhebung im Original.

⁸⁸ Vgl. Schermelleh-Engel/Werner 2007, S. 114.

⁸⁹ Für die Herleitung der Formel vgl. Backhaus et alii 2006, S. 377 f.

auf die Mindestreliabilität für die Indikatoren gibt es in der Literatur verschiedene Aussagen. Homburg/Baumgartner geben einen Zielwert von 0,4 bis 0,5 an,⁹⁰ Nunnally empfiehlt einen Schwellenwert von 0,5 bis 0,6.⁹¹ Hair/Bush/Ortinou empfehlen Reliabilitätswerte von mindestens 0,6,⁹² und Nunnally/Bernstein empfehlen einen Grenzwert von 0,7.⁹³ Inhaltlich bedeutet dies, dass ein Indikator als reliabel betrachtet wird, wenn seine Varianz zu 40 Prozent (bzw. 50, 60 oder 70 Prozent) oder mehr durch den Faktor erklärt werden kann, den der Indikator repräsentiert. Welche Indikatorreliabilität einen angemessenen Grenzwert darstellt, hängt auch vom Stichprobenumfang ab.⁹⁴ Für einen Stichprobenumfang zwischen 400 und 1000, wie in der vorliegenden Studie, empfiehlt Balderjahn Untergrenzen für die Indikatorreliabilität von 0,2 bis 0,4.⁹⁵ Im Folgenden wird deshalb ein Anspruchsniveau von $\geq 0,4$ für die Indikatorreliabilität verwendet.

Der t-Wert der Faktorladungen wird durch den Quotient von Faktorladung und Standardfehler gebildet. Für einen zweiseitigen Signifikanztest entspricht 2,33 einem Signifikanzniveau von 0,01 (als Näherungswert bei einer großen Anzahl von Freiheitsgraden).⁹⁶ Dieser Wert wird für die vorliegende Arbeit als Anspruchsniveau gewählt.

(2) Die Reliabilität eines *Faktors* gibt an, inwiefern die Items das gleiche unidimensionale Konstrukt messen.⁹⁷ Entsprechend der Definition der Reliabilität kann die Faktorreliabilität als Grad aufgefasst werden, zu dem ein Konstrukt frei von *zufälligen* Fehlern ist.⁹⁸ Je höher die Reliabilität eines Faktors, desto geringer ist der Messfehler und desto ähnlichere Ergebnisse sind zu erwarten, wenn man das gleiche Konstrukt mehrmals im Verlauf der Zeit abfragt.⁹⁹ Für die Beurteilung der Reliabilität eines Faktors stellt Cronbachs α ein häufig verwendetes Maß dar.¹⁰⁰ Auch dieser Wert kann zwischen null und eins liegen, wobei höhere Werte eine höhere Reliabilität widerspiegeln. Auch für die erforderliche Faktorreliabilität (construct reliability) gibt es in der Literatur verschiedene Aussagen. So fordern Bagozzi/Yi einen Wert von größer oder

⁹⁰ Vgl. Homburg/Baumgartner 1995, S. 170.

⁹¹ Vgl. Nunnally 1967, S. 226.

⁹² Vgl. Hair/Bush/Ortinou 2000, S. 391.

⁹³ Vgl. Nunnally/Bernstein 1994, S. 265.

⁹⁴ Vgl. Balderjahn 1986, S. 116 f.

⁹⁵ Vgl. Balderjahn 1986, S. 117.

⁹⁶ Vgl. Backhaus et alii 2006, S. 808.

⁹⁷ Vgl. Dunn/Seaker/Waller 1994.

⁹⁸ Vgl. Hildebrandt/Temme 2006, S. 621.

⁹⁹ Vgl. Garver/Mentzer 1999, S. 35.

¹⁰⁰ Vgl. Cronbach 1951.

gleich 0,6,¹⁰¹ Garver/Mentzer hingegen einen Wert von größer oder gleich 0,7.¹⁰² Die Höhe von Cronbachs α wird von der durchschnittlichen Korrelation der Indikatoren beeinflusst. Für die vorliegende Arbeit wird eine Faktorreliabilität von $\geq 0,6$ als zufriedenstellend erachtet.

(3) Wie stark ein Indikator mit allen anderen Indikatoren korreliert, wird von der sogenannten korrigierten Item-to-Total-Korrelation jedes Indikators angegeben. Darunter versteht man die Korrelation eines Indikators mit der Summe aller übrigen Indikatoren des entsprechenden Faktors. Items mit einer hohen Item-to-Total-Korrelation verbessern die Reliabilität des Faktors.¹⁰³ Je höher die Item-to-Total-Korrelation jedes Indikators, desto höher ist demzufolge auch Cronbachs α . Gleichzeitig kann Cronbachs α gesteigert werden, wenn Indikatoren mit besonders geringen Item-to-Total-Korrelationen eliminiert werden.¹⁰⁴ Für die Item-to-Total-Korrelation existiert kein fester Grenzwert.

Im Kapitel 4.3 auf Seite 121 ff. werden für jede Skala Cronbachs α und für jeden Indikator eines latenten Konstrukts die Indikatorreliabilität, der t-Wert der Faktorladung und die korrigierte Item-to-Total-Korrelation aufgeführt.

4.2.2.2 Kriterien zur Prüfung der Unidimensionalität

Ein wichtiges Merkmal reflektiver Konstrukte ist die Unidimensionalität. Sie kann definiert werden als die Existenz genau eines latenten Merkmals für ein Set von Indikatoren.¹⁰⁵ Anders ausgedrückt bedeutet Unidimensionalität also, dass alle Indikatoren eines Konstrukts *genau eine* latente Variable repräsentieren.¹⁰⁶ Der Unidimensionalität kommt eine noch größere Bedeutung zu als der Reliabilität: „Auch wenn hohe Reliabilitäten für die Analyse prinzipiell immer wünschenswert sind, ist aber entscheidender, dass die Messmodelle wirklich *eindimensional* sind, also keine Korrelationen zwischen den Messfehlern vorliegen.“¹⁰⁷

Um die Unidimensionalität der Messmodelle zu prüfen, wird in der Literatur der Einsatz der *konfirmatorischen* Faktorenanalyse empfohlen.¹⁰⁸ Wenn das Messmodell eindimensional ist,

¹⁰¹ Vgl. Bagozzi/Yi 1988, S. 82.

¹⁰² Vgl. Garver/Mentzer 1999, S. 44.

¹⁰³ Vgl. Nunnally 1978, S. 279 f.

¹⁰⁴ Vgl. Churchill 1979, S. 68.

¹⁰⁵ Vgl. Anderson/Gerbing/Hunter 1987, S. 432.

¹⁰⁶ Vgl. Garver/Mentzer 1999, S. 35.

¹⁰⁷ Scholderer/Balderjahn/Paulssen 2006, S. 646; Hervorhebung im Original.

¹⁰⁸ Vgl. beispielsweise Anderson/Gerbing/Hunter 1987, S. 435, und Gerbing/Anderson 1988, S. 186 f.

liegen keine Korrelationen zwischen den Messfehlern vor. Denn die Korrelationen sind ein Indiz dafür, dass die Items neben der latenten Variablen, die sie repräsentieren sollen, noch eine weitere gemeinsame Dimension besitzen.¹⁰⁹ Die Ergebnisse der konfirmatorischen Faktorenanalyse werden in Abschnitt 4.3.5.1 auf Seite 144 ff. dargestellt.

4.2.2.3 Kriterien zur Prüfung der Validität

„*Validität* ist ein integriertes bewertendes Urteil über das Ausmaß, in dem die Angemessenheit und Güte von Interpretationen und Maßnahmen auf Basis von Testwerten und anderen diagnostischen Verfahren durch empirische Belege und theoretische Argumente gestützt sind.“¹¹⁰ Die Konstruktvalidität drückt demnach aus, inwieweit ein Konstrukt tatsächlich das misst, was der Forscher damit zu messen beabsichtigt.¹¹¹ Die Validität ist umso höher, je geringer der Zufallsfehler *und* der systematische Fehler einer Messung sind. In der Literatur werden zahlreiche Dimensionen der Validität diskutiert.¹¹² Im Folgenden sollen die Aspekte (1) Inhaltsvalidität, (2) Konvergenzvalidität, (3) Diskriminanzvalidität und (4) externe Validität näher erläutert werden.

(1) Die *Inhaltsvalidität* (Content Validity) beschreibt das Ausmaß, indem das Konstrukt durch die Items repräsentiert wird.¹¹³ Für die Inhaltsvalidität existiert kein formaler statistischer Test.¹¹⁴ Umso wichtiger ist es deshalb, schon bei der Entwicklung eines Konstrukts auf Inhaltsvalidität zu achten.¹¹⁵ Für die vorliegende Studie wurde dies durch einen klaren Prozess der Skalenentwicklung¹¹⁶ und mehrere Pretests mit Personen aus der Zielpopulation sichergestellt.¹¹⁷

(2) Der Begriff *Konvergenzvalidität* gibt an, wie stark verschiedene Indikatoren das gleiche Konstrukt widerspiegeln. Einen Anhaltspunkt für die Konvergenzvalidität gibt die exploratori-

¹⁰⁹ Vgl. Scholderer/Balderjahn/Paulssen 2006, S. 646.

¹¹⁰ Messick 1989, S. 13, zitiert in der Übersetzung von Hartig/Frey/Jude 2007, S. 136.

¹¹¹ Vgl. Garver/Mentzer 1999, S. 34.

¹¹² So unterscheiden beispielsweise Garver/Mentzer sieben Aspekte der Konstruktvalidität: Inhaltsvalidität, substantielle Validität, Unidimensionalität, Reliabilität, Konvergenzvalidität, Diskriminanzvalidität und vorhersagende Validität; vgl. Garver/Mentzer 1999, S. 33.

¹¹³ Vgl. Garver/Mentzer 1999, S. 34.

¹¹⁴ Vgl. Garver/Mentzer 1999, S. 34.

¹¹⁵ So konstatiert beispielsweise Hensley: “The ultimate purpose of item generation is to ensure that questionnaire items have content validity.”; Hensley 1999, S. 348.

¹¹⁶ Vgl. Abschnitt 3.2.2 auf Seite 52.

¹¹⁷ Vgl. Abschnitt 3.4.1 auf Seite 76.

sche Faktorenanalyse. Der erste extrahierte Faktor sollte gemäß der Empfehlungen von Homburg/Giering mindestens 50 Prozent der Varianz der ihm zugeordneten Indikatoren erklären.¹¹⁸ Diesem Richtwert schließt sich die vorliegende Arbeit an. Die entsprechende Größe wird im Folgenden als „Varianzerklärung durch den ersten Faktor“ bezeichnet und für jedes latente Konstrukt angegeben.¹¹⁹

Die statistische Prüfung der Konvergenzvalidität kann außerdem mit Hilfe der konfirmatorischen Faktorenanalyse erfolgen.¹²⁰ Dabei wird die durchschnittlich erfasste Varianz (DEV) ermittelt. Diese drückt aus, welcher Anteil der Varianz der Indikatoren durchschnittlich von dem übergeordneten Faktor erklärt wird.¹²¹ Die DEV kann Werte zwischen null und eins annehmen, wobei ein höherer Wert eine höhere Konvergenzvalidität ausdrückt. Als Mindestwert wird üblicherweise 0,5 verlangt.¹²² Diesem Anspruch folgt die vorliegende Arbeit. Inhaltlich bedeutet dies, dass der Faktor einen größeren Anteil der Varianz der Indikatoren bedingt als der Messfehler.¹²³ Im Kapitel 4.3 auf Seite 121 ff. wird für jedes latente Konstrukt die DEV angegeben.

(3) Der Begriff *Diskriminanzvalidität* steht für die Forderung, dass Messungen desselben Merkmals (beispielsweise Indikatoren eines latenten Konstrukts) stärker miteinander korrelieren sollten als Messungen verschiedener Merkmale (beispielsweise Indikatoren unterschiedlicher Konstrukte).¹²⁴ Die Prüfung der Diskriminanzvalidität kann ebenso wie die Konvergenzvalidität im Rahmen einer konfirmatorischen Faktorenanalyse erfolgen.¹²⁵ Das Fornell/Larcker-Kriterium stellt dabei ein verbreitetes Maß zur Bewertung der Diskriminanzvalidität dar.¹²⁶ Es besagt, dass Diskriminanzvalidität vorliegt, wenn die quadrierten Korrelationen zwischen zwei Konstrukten kleiner sind als die durchschnittlich erfassten Varianzen jedes der beiden Faktoren.¹²⁷ Die Diskriminanzvalidität aller Konstrukte wird ausführlich in Abschnitt 4.3.5 auf Seite 143 ff. analysiert.

¹¹⁸ Vgl. Homburg/Giering 1996, S. 12.

¹¹⁹ Vgl. Homburg/Giering 1996, S. 16.

¹²⁰ Vgl. Hildebrandt/Temme 2006, S. 621.

¹²¹ Vgl. Fornell/Larcker 1981, S. 45.

¹²² Vgl. Homburg/Baumgartner 1995, S. 170.

¹²³ Vgl. Vandebosch 1996, S. 23 ff.

¹²⁴ Vgl. Campbell/Fiske 1959.

¹²⁵ Vgl. Hildebrandt/Temme 2006, S. 621.

¹²⁶ Vgl. Fornell/Larcker 1981.

¹²⁷ Vgl. Fornell/Larcker 1981, S. 39 ff.

(4) Die *externe Validität* adressiert die Fragestellung, inwieweit Forschungsergebnisse auf eine breitere Population übertragen werden können.¹²⁸ Die externe Validität wird maßgeblich durch die Auswahl der Befragten und die untersuchte Situation bestimmt. Da die vorliegende Studie Situationen aus dem realen Unternehmenskontext branchenübergreifend untersucht, ist von einer hohen externen Validität auszugehen.¹²⁹

Tabelle 4.3: Gütebeurteilung von Messmodellen

Anpassungsmaß	Anspruchsniveau
Indikatorreliabilität	$\geq 0,40$
t-Wert der Faktorladung	$\geq 2,33$
Faktorreliabilität	$\geq 0,60$
Cronbachs α	$\geq 0,60$
Varianzerklärung durch den ersten Faktor	$\geq 0,50$
Item-to-Total-Korrelation	$< 0,08$
DEV	$\geq 0,50$

Tabelle 4.3 fasst die bislang diskutierten Anspruchsniveaus zusammen. Umfasst ein Messmodell vier oder mehr Indikatoren, können weitere Anpassungsmaße für die Beurteilung herangezogen werden. Dabei handelt es sich um die gleichen Gütemaße, die für die Beurteilung von Strukturmodellen verwendet werden. Sie werden in Abschnitt 4.2.3 vorgestellt.

4.2.3 Kriterien für die Gütebeurteilung der Mess- und Strukturmodelle

Zur Beurteilung von Strukturgleichungsmodellen (sowie Messmodellen mit mindestens vier Indikatoren) wird das Ausmaß der Übereinstimmung zwischen der modellimplizierten und der empirischen Kovarianzmatrix geprüft.¹³⁰ Dies erfolgt mit Hilfe sogenannter Güte- oder Anpassungsmaße.¹³¹ Das Spektrum von in der Literatur vorgeschlagenen Anpassungsmaßen ist sehr

¹²⁸ Vgl. Garver/Mentzer 1999, S. 34.

¹²⁹ Für die Auswahl der Informanten vgl. Abschnitt 3.1.2 auf Seite 48, für deskriptive Auswertungen Abschnitt 4.1.3 auf Seite 107.

¹³⁰ Vgl. Schermelleh-Engel/Moosbrugger/Müller 2003, S. 23.

¹³¹ „Anpassungsmaße sind Größen, die auf der Basis der Parameterschätzung die Güte der Anpassung des relevanten Modells an den vorliegenden Datensatz beurteilen.“; Homburg/Baumgartner 1998, S. 351; Hervorhebung im Original.

groß, weist jedoch gleichzeitig weitreichende Redundanzen auf.¹³² Deshalb ist es nicht notwendig, alle Beurteilungsgrößen zu betrachten.¹³³ Stattdessen ist eine *ausgewogene* Auswahl von Größen sinnvoll, wie sie beispielsweise von Garver/Mentzer,¹³⁴ Homburg/Baumgartner¹³⁵ und Homburg/Klarmann¹³⁶ empfohlen wird.

Als Kriterien für die Auswahl der Indizes gelten die relative Unabhängigkeit vom Stichprobenumfang, die Genauigkeit und Konsistenz bei der Beurteilung verschiedener Modelle sowie die einfache Interpretation durch einen definierten Ergebnisbereich.¹³⁷ Homburg/Klarmann empfehlen auf der Basis einer Reihe von Studien, insbesondere folgende Gütekriterien zu prüfen:¹³⁸ (1) $\frac{\chi^2}{df}$, (2) TLI (= NNFI), (3) CFI, (4) RMSEA und (5) SRMR. Dieser Empfehlung wird in der vorliegenden Studie entsprochen.¹³⁹ Die Anpassungsmaße werden im Folgenden näher erläutert. Für jedes Maß ist in der Literatur eine Reihe von Empfehlungen zu finden, welcher Grenzwert einen akzeptablen Fit darstellt.¹⁴⁰ In den folgenden Beschreibungen der Anpassungsmaße wird jeweils ein üblicher Wert angegeben, der als Richtwert ohne Anspruch auf Exaktheit aufzufassen ist.

(1) Das Anpassungsmaß $\frac{\chi^2}{df}$ setzt die χ^2 -Teststatistik in Relation zu den Freiheitsgraden (df). Die χ^2 -Teststatistik ist ein Maß dafür, wie stark die modellinduzierte Kovarianzmatrix von der empirischen Matrix abweicht. Liegt der Wert von $\frac{\chi^2}{df}$ zwischen zwei und drei, gilt der Fit als gut bis akzeptabel.¹⁴¹ In anderen Worten: Die χ^2 -Teststatistik sollte maximal dreimal so hoch sein wie die Anzahl der Freiheitsgrade. Für die vorliegende Arbeit wird deshalb im Folgenden ein Wert von $\frac{\chi^2}{df}$ von kleiner als drei als Anspruchsniveau verwendet.

¹³² Für eine kurze Erläuterung der Klassifikation von Anpassungsmaßen vgl. Homburg 1992b, S. 504 ff., und Schermelleh-Engel/Moosbrugger/Müller 2003, S. 31 ff.

¹³³ Vgl. beispielsweise Homburg/Baumgartner 1998, S. 369.

¹³⁴ Vgl. Garver/Mentzer 1999, S. 41.

¹³⁵ Vgl. Homburg/Baumgartner 1998, S. 363.

¹³⁶ Vgl. Homburg/Klarmann 2006, S. 736.

¹³⁷ Vgl. Marsh/Balla/McDonald 1988.

¹³⁸ Vgl. Homburg/Klarmann 2006, S. 736.

¹³⁹ Die weit verbreiteten Anpassungsmaße GFI und AGFI gelten in der aktuellen Literatur nicht mehr als beste Wahl und werden deshalb nicht verwendet; vgl. Homburg/Klarmann 2006, S. 736.

¹⁴⁰ Vgl. beispielsweise Garver/Mentzer 1999, Homburg 1992b, Homburg/Baumgartner 1995, Homburg/Giering 1996, Homburg/Klarmann 2006, Hu/Bentler 1995, Hu/Bentler 1998, Hu/Bentler 1999 und Schermelleh-Engel/Moosbrugger/Müller 2003.

¹⁴¹ Vgl. Homburg/Baumgartner 1995, Homburg/Giering 1996, Homburg/Klarmann 2006 und Schäffer 2008, S. 7.

(2) Der Tucker-Lewis-Index (TLI), der auch als Nonnormed-Fit-Index (NNFI) bezeichnet wird, vergleicht den Fit des Modells mit einem Basismodell bzw. Unabhängigkeitsmodell.¹⁴² Zusätzlich berücksichtigt der TLI die Freiheitsgrade des vorgeschlagenen Modells und des Basismodells. Der TLI ist unabhängig von der Stichprobengröße und wird in der Literatur nachdrücklich empfohlen.¹⁴³ Der TLI kann zwischen null und eins liegen, Werte von 0,95 und höher werden als akzeptabel angesehen.¹⁴⁴ Werte von über 0,97 weisen auf einen guten Fit hin.¹⁴⁵ Für die vorliegende Arbeit kommt deshalb im Folgenden ein Anspruchsniveau von TLI größer als 0,95 zum Einsatz.

(3) Der Comparative-Fit-Index (CFI) ist ebenfalls ein von der Stichprobengröße unabhängiger Index zur Bewertung der Modellgüte.¹⁴⁶ Der CFI kann Werte von null bis eins einnehmen, wobei höhere Werte einen besseren Fit anzeigen. Werte von 0,95 und höher werden als akzeptabel angesehen.¹⁴⁷ Werte über 0,97 gelten als gut.¹⁴⁸ Für die vorliegende Arbeit wird ein Anspruchsniveau von CFI größer 0,95 gewählt.

(4) Der Root Mean Square Error of Approximation (RMSEA)¹⁴⁹ prüft den approximativen Fit der Population.¹⁵⁰ Je näher der RMSEA an null liegt, desto besser der Fit. Werte bis 0,05 gelten als gut, Werte kleiner 0,08 gelten als zufriedenstellend,¹⁵¹ Werte bis 0,10 meist als gerade noch akzeptabel.¹⁵² Für die vorliegende Arbeit wird ein Anspruchsniveau von RMSEA kleiner 0,08 als zufriedenstellend betrachtet.

(5) Der Standardized Root Mean Square Residual (SRMR)¹⁵³ bezieht den Fit auf die standardisierten Residuen. Ein Wert von null entspricht perfektem Fit. Werte kleiner als 0,05 indizieren

¹⁴² Vgl. Tucker/Lewis 1973.

¹⁴³ Vgl. Garver/Mentzer 1999, S. 41.

¹⁴⁴ Von Hu/Bentler werden 0,95 als gut eingestuft; vgl. Hu/Bentler 1999.

¹⁴⁵ Vgl. Schermelleh-Engel/Moosbrugger/Müller 2003, S. 41. Andere Autoren halten auch 0,9 für akzeptabel; vgl. Garver/Mentzer 1999, S. 41, Homburg 1992b und Homburg/Klarmann 2006.

¹⁴⁶ Vgl. Bentler 1990.

¹⁴⁷ Von Hu/Bentler werden 0,95 als gut eingestuft; vgl. Hu/Bentler 1999.

¹⁴⁸ Vgl. Schermelleh-Engel/Moosbrugger/Müller 2003, S. 42 und Schäffer 2008, S. 10. Andere Autoren halten auch 0,9 für akzeptabel; vgl. auch Garver/Mentzer 1999, S. 41, Homburg/Baumgartner 1995 und Homburg/Klarmann 2006.

¹⁴⁹ Vgl. Steiger/Lind 1980 und Steiger 1990.

¹⁵⁰ Vgl. Schermelleh-Engel/Moosbrugger/Müller 2003, S. 36.

¹⁵¹ Vgl. Browne/Cudeck 1993, S. 136 ff., und Garver/Mentzer 1999, S. 41.

¹⁵² Vgl. Browne/Cudeck 1993, Homburg/Baumgartner 1995, Homburg/Klarmann 2006 und Hu/Bentler 1999.

¹⁵³ Vgl. Jöreskog/Sörbom 1981, S. 41 und Jöreskog/Sörbom 1989.

einen guten Fit, Werte bis 0,10 gelten als akzeptabel.¹⁵⁴ Für die vorliegende Arbeit wird ein SRMR kleiner 0,08 als zufriedenstellend erachtet.

Tabelle 4.4: Gütebeurteilung von Strukturmodellen

Anpassungsmaß	Anspruchsniveau
$\frac{\chi^2}{df}$	< 3,00
TLI	> 0,95
CFI	> 0,95
RMSEA	< 0,08
SRMR	< 0,08

Die Anspruchsniveaus für die Gütebeurteilung der Strukturmodelle sind in Tabelle 4.4 zusammengefasst. Die Werte der Anpassungsmaße werden im Folgenden in jeder Abbildung von Strukturgleichungsmodellen angegeben. Die variierenden Empfehlungen aus der vorgestellten Literatur zeigen, dass es keine exakten, eindeutigen Grenzwerte für die Beurteilung von Strukturmodellen gibt. Die Anspruchsniveaus stellen also lediglich Richtwerte dar und es ist *nicht* notwendig, dass ein Modell für *alle* Anpassungsmaße die Anspruchsniveaus erfüllt.¹⁵⁵ Auch wenn einzelne Anspruchsniveaus leicht verfehlt werden, kann ein Modell akzeptiert werden, wenn die Gütekriterien insgesamt einen positiven Gesamteindruck vermitteln.¹⁵⁶ Gleichzeitig bietet das Erfüllen aller Anspruchsniveaus keine Gewähr dafür, dass ein Modell inhaltlich sinnvoll ist. Vielmehr muss die Angemessenheit eines Modells stets unter Berücksichtigung von theoretischen, statistischen und praktischen Aspekten beurteilt werden.¹⁵⁷

¹⁵⁴ Vgl. Homburg/Klarmann 2006 und Schermelleh-Engel/Moosbrugger/Müller 2003, S. 38.

¹⁵⁵ Vgl. Homburg/Baumgartner 1998, S. 363.

¹⁵⁶ Vgl. Homburg/Pflesser 2000, S. 655.

¹⁵⁷ Vgl. Byrne 2001, S. 88.

4.3 Entwicklung und Beurteilung der Messmodelle

Hildebrandt/Temme 2006 weisen darauf hin, dass der Prozess der Skalenoptimierung *nicht* mechanisch durchgeführt werden sollte. So kritisieren sie beispielsweise stellvertretend für viele Artikel den Aufsatz von Ewing/Napoli,¹⁵⁸ in dem ohne Diskussion fast 50 Prozent der Indikatoren ausschließlich anhand der korrigierten Item-to-Total-Korrelation eliminiert werden.¹⁵⁹ Um dieser Kritik zu begegnen, wird im weiteren Verlauf neben der Beurteilung der statistischen Größen jeweils auch sachlogisch analysiert, warum ein Indikator sich als nicht reliabel erweist.

Die Gütekriterien der Messmodelle werden jeweils mit dem gesamten Datensatz berechnet. Dies basiert auf der Annahme, dass die Messmodelle für Entscheider und Nicht-Entscheider gleich sind, auch wenn die Strukturmodelle Unterschiede aufweisen können.

4.3.1 Operationalisierung der Projekteskalation

Unter Projekteskalation wird der ökonomisch-irrationale, systematisch verspätete Abbruch eines fehlgeschlagenen Projekts verstanden.¹⁶⁰ Die Projekteskalation ist umso größer, desto mehr Ressourcen (beispielsweise Kosten, Mannstunden und Kalenderzeit) seit dem Zeitpunkt in das Projekt geflossen sind, ab dem aus ökonomisch-rationaler Sicht das Projekt hätte abgebrochen werden müssen. Weil davon auszugehen ist, dass dieser Zeitpunkt den Befragten nicht exakt bekannt und eine objektive Messung im Nachhinein kaum möglich ist,¹⁶¹ soll versucht werden, das Konstrukt mit verschiedenen Verfahren näherungsweise zu messen. Bevor die angewandte Messung für die Projekteskalation vorgestellt wird, werden kurz die bisher in der Literatur genannten Messungen diskutiert.

Armstrong/Williams/Barrett verwenden als Messung für die Eskalationstendenz einen Index, den sie aus der Verweildauer des Top-Managements im Unternehmen und den Veränderungen der Erträge berechnen.¹⁶² Eine Eskalation liegt demnach dann vor, wenn ein Manager trotz

¹⁵⁸ Vgl. Ewing/Napoli 2005.

¹⁵⁹ Vgl. Hildebrandt/Temme 2006, S. 619.

¹⁶⁰ Vgl. hierzu auch die Definition in Tabelle 3.1 auf Seite 54.

¹⁶¹ So stellen Armstrong/Williams/Barrett fest: "The most difficult variable to measure after *unethical behavior* is *escalating commitment*."; Armstrong/Williams/Barrett 2004, S. 368; Hervorhebungen im Original.

¹⁶² Vgl. Armstrong/Williams/Barrett 2004, S. 368.

schlechter Performance seines Zuständigkeitsbereichs seine Position behält. Die Gültigkeit dieser Messung ist jedoch fragwürdig, wie Armstrong/Williams/Barrett selbst einräumen,¹⁶³ und außerdem nicht im Projektbereich anwendbar, da die Zeiträume kürzer sind und nicht davon ausgegangen werden kann, dass bereits eine Eskalation des Commitments vorliegt, wenn ein Projektmanager, dessen Projekt schlecht läuft, nicht gekündigt wird.

In einer Studie von Keil et alii 2003 wird zwischen Projekten, die eskaliert sind, und solchen, bei denen dies nicht der Fall ist, unterschieden.¹⁶⁴ Die Variable Projekteskalation wird also dichotom gemessen. Die jeweilige Klassifizierung wurde durch eine einzelne Frage an die Teilnehmer festgelegt: "Participants who received the 'escalation' version of the survey were asked to identify a *troubled* project that continued to receive resources even though s/he thought that the project should have been discontinued or redirected."¹⁶⁵

Keil et alii 2000 verwenden in einer Studie zur Eskalation des Commitments die Variable „Willingness to continue a project“ als Indikator für die Eskalation. Die Variable wird dabei mit nur einem Item gemessen: "How likely is it that you personally would choose to continue with the [...] project?", wobei die Antwortskala von „0 % = definitely would not continue“ bis „100 % = definitely would continue“ reicht.¹⁶⁶ Auch diese Messung ist nicht für die Ex-post-Betrachtung der vorliegenden Studie geeignet.

Keine der genannten Operationalisierungen kann überzeugen, da es sich entweder um Single Items handelt und/oder die Messung für die Eskalation risikoreicher Projekte ungeeignet ist. Deshalb wird in der vorliegenden Studie der Versuch unternommen, die Projekteskalation differenzierter zu messen. Dazu werden mehrere Konstrukte mit jeweils mehreren Indikatoren verwendet.

Ein Konstrukt zur Messung der Projekteskalation trägt die Bezeichnung „verspäteter Projektabbruch“. Hierbei werden reflektive Indikatoren gewählt, um die *subjektive Einschätzung* der Befragten, ob das Projekt zu spät abgebrochen wurde, zu ermitteln. Eine Aussage dazu lautet beispielsweise: „Im Nachhinein betrachtet, wurde das Projekt zu spät abgebrochen.“¹⁶⁷

Um weitere Anregungen für mögliche Konstrukte zu gewinnen, kann ein Blick in die Literatur zur Messung des Projekterfolgs dienen. In der Projektmanagementliteratur werden zur Messung

¹⁶³ Vgl. Armstrong/Williams/Barrett 2004, S. 369.

¹⁶⁴ Vgl. Keil et alii 2003.

¹⁶⁵ Keil et alii 2003, S. 245; Hervorhebung im Original.

¹⁶⁶ Vgl. Keil et alii 2000, S. 308.

¹⁶⁷ Für den Wortlaut der Indikatoren und die statistischen Gütemaße vgl. Abschnitt 4.3.1.1 auf Seite 124.

des Projekterfolgs vor allem die drei Kriterien Budget, Zeitplan und Performance sowie gelegentlich als viertes Kriterium die Kundenzufriedenheit genannt.¹⁶⁸ Bei der Untersuchung *abgebrochener* bzw. *gescheiterter* Projekte stellt sich die Frage, welche Elemente des Projekterfolgs zur Messung der Eskalation dieser Projekte dienen können. Passend für die Messung der Eskalation erscheinen das Budget und der Zeitplan.¹⁶⁹ Dementsprechend soll versucht werden, das Ausmaß der Eskalation über den Ressourcenverbrauch (Budget, Zeit) bis zum Projektabbruch zu ermitteln. Dazu werden zwei Konstrukte verwendet.

(1) Ein Konstrukt fragt nach der *Beurteilung* der eingesetzten Ressourcen (Manntage, Kosten und Kalendermonate). Dies deckt sich mit dem Verständnis der Eskalation des Commitments in der Literatur.¹⁷⁰ Eskalation bezieht sich demnach nicht nur auf einen speziellen Faktor, sondern allgemein auf die wachsende, kumulative Menge von investierten Ressourcen im Zeitverlauf.¹⁷¹ Das Konstrukt trägt die Bezeichnung „Beurteilung des Ressourcen-Einsatzes“ und wird wie die folgenden Konstrukte in einem eigenen Abschnitt vorgestellt.¹⁷²

(2) Ein zweites Konstrukt bezieht sich auf dieselben Dimensionen, fragt jedoch nach dem *objektiven Ressourcenverbrauch*. Die Formulierung lautet: „Welcher Anteil der ursprünglich (bei Projektbeginn) geplanten Ressourcen wurde bis zum Abbruch des Projekts eingesetzt?“¹⁷³ Dieses Konstrukt wird im Folgenden als „investierter Anteil der geplanten Ressourcen“ bezeichnet. Das Konstrukt wird in Abschnitt 4.3.1.3 näher vorgestellt.

Die drei erläuterten Konstrukte bilden zusammen die Messung für die Projekteskalation, die deshalb als Konstrukt zweiter Ordnung operationalisiert wird.¹⁷⁴

¹⁶⁸ Vgl. exemplarisch Pinto/Slevin 1988, S. 68. Ähnlich Ashley/Lurie/Jaselskis, die statt „Performance“ den Begriff „Functionality“ verwenden und die Zufriedenheit auch auf den Kontraktor, den Projektmanager und das Team ausweiten; vgl. Ashley/Lurie/Jaselskis 1987, S. 69. Für eine detailliertere Messung des Projekterfolgs vgl. Dvir/Raz/Shenhar 2003, S. 92. Für eine Messung von Produktentwicklungs- und Managementperformance vgl. Kahn 1996, S. 143.

¹⁶⁹ Die Kundenzufriedenheit scheidet klar aus, da bei einem abgebrochen bzw. gescheiterten Projekt in der Regel keine Kundenzufriedenheit im Sinne eines Projekterfolgs vorliegen kann. Die Performance kann ebenfalls nicht als Kriterium herangezogen werden, weil das Ziel des Projekts (Produkt, Dienstleistung oder sonstiges Resultat; vgl. die Definition von Projekten in Abschnitt 2.1.2 auf Seite 12) in der Regel nicht erreicht wird und selbst bei erfolgreicher Erstellung nicht im geplanten Sinn eingesetzt wird.

¹⁷⁰ Vgl. Keil et alii 1995, S. 80, Newman/Sabherwal 1996, S. 24, und Whyte 1986.

¹⁷¹ Vgl. Keil et alii 1995, S. 80.

¹⁷² Vgl. Abschnitt 4.3.1.2. Der Wortlaut der Indikatoren und die statistischen Gütemaße sind ebenfalls in Abschnitt 4.3.1.2 zu finden.

¹⁷³ Für den Wortlaut der Indikatoren und die statistischen Gütemaße vgl. Abschnitt 4.3.1.3 auf Seite 126.

¹⁷⁴ Vgl. Abschnitt 4.3.1.4 auf Seite 127.

4.3.1.1 Verspätung des Projektabbruchs

Mit dem Konstrukt verspäteter Projektabbruch soll die empfundene Verzögerung des Projektabbruchs gemessen werden, im Vergleich zu dem Zeitpunkt, ab dem rational betrachtet das Projekt hätte abgebrochen werden müssen.¹⁷⁵ Die Antwortmöglichkeiten liegen auf einer siebenstufigen Skala von „0 = trifft gar nicht zu“ bis „6 = trifft voll zu“. Der Wortlaut der Fragen ist in Tabelle 4.5 angegeben.

Tabelle 4.5: Gütekriterien des Messmodells „verspäteter Projektabbruch“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Im Nachhinein betrachtet, ...				
1.	... wurde das Projekt zu spät abgebrochen.	0,87	0,86	fixiert
2.	... wurde zu hartnäckig an dem Projekt festgehalten.	0,81	0,71	31,46
3.	... hätte das Projekt früher abgebrochen werden müssen.	0,85	0,83	36,19
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,92, Varianzklärung durch den ersten Faktor: 0,87, Faktorreliabilität: 0,92 und DEV: 0,80				

Tabelle 4.5 stellt außerdem die Gütekriterien für die Einzelindikatoren dar. Alle Items erfüllen die Anspruchsniveaus.¹⁷⁶ Die Indikatorreliabilitäten sind größer als 0,4 und die t-Werte der Faktorladungen sind größer als 2,33. Auch das Gesamtkonstrukt weist zufriedenstellende Gütemaße auf. Alle Werte liegen über den geforderten Mindestgrenzen: Cronbachs α über 0,7, die Varianzklärung durch den ersten Faktor über 0,5, die Faktorreliabilität über 0,6 und die DEV über 0,5. Alle Werte sprechen somit für eine hohe Güte des Messmodells. Eine Modifikation ist nicht notwendig.

Die Anpassungsmaße $\frac{\chi^2}{df}$, TLI, CFI, RMSEA und SRMR können erst ab einer Mindestanzahl von vier Indikatoren ermittelt werden, da ein Messmodell mit drei Indikatoren keine Freiheitsgrade aufweist. Diese Anpassungsmaße werden deshalb hier und bei allen folgenden Konstrukten mit drei oder weniger Items *nicht* angegeben.

¹⁷⁵ Es handelt sich also um eine subjektive Einschätzung der „tendency to continue uneconomic projects“; Cheng et alii 2003, S. 64.

¹⁷⁶ Vgl. Abschnitt 4.2.2 auf Seite 112 ff.

4.3.1.2 Beurteilung des Ressourcen-Einsatzes

Für das Konstrukt Beurteilung des Ressourcen-Einsatzes wurden die Befragten gebeten, die Einsatzmenge der Ressourcen zu bewerten. Dies wurde im Hinblick auf drei Dimensionen abgefragt: Mannmonate, Kosten und Dauer. Die Antworten erfolgten auf einer siebenstufigen Skala von „-3 (viel zu wenig)“ bis „+3 (viel zu viel)“.

Tabelle 4.6: Gütekriterien des Messmodells „Beurteilung des Ressourcen-Einsatzes“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Bitte bewerten Sie, ob zu wenig, genau richtig oder zu viele Ressourcen in das abgebrochene Projekt geflossen sind.				
1.	Mannmonate	0,63	0,62	fixiert
2.	Kosten (finanzielle Ressourcen)	0,60	0,53	13,87
3.	Dauer des Projekts (Kalenderzeit)	0,56	0,43	13,53
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,76, Varianzerklärung durch den ersten Faktor: 0,68, Faktorreliabilität: 0,81 und DEV: 0,53				

Die Gütemaße des Messmodells erreichen teilweise nur knapp die geforderten Werte. Insbesondere die Indikatorreliabilität der Dauer des Projekts ist mit 0,43 nahe am zulässigen Grenzwert. Auch die DEV liegt nur knapp über dem Schwellenwert von 0,5. Inhaltlich bedeutet dies, dass die Indikatoren neben dem gemeinsamen Faktor auch verhältnismäßig stark von Faktoren beeinflusst werden, die sich zwischen den Indikatoren unterscheiden. Dies ist durchaus plausibel. Beispielsweise kam in Interviews mit Praktikern zur Sprache, dass scheiternde Projekte häufig formal betrachtet noch eine Zeitlang weiterlaufen, jedoch keine Ressourcen mehr investiert werden. In diesen Fällen ist Dauer des Projektes (Kalenderzeit) von den investierten Mannmonaten und Kosten abgekoppelt.

Die Anspruchsniveaus werden nicht unterschritten, was dafür spricht, dass die Indikatoren trotz ihrer verschiedenen Bezugsgrößen einen ausreichenden Anteil eines gemeinsamen Faktors abdecken. Da zudem alle drei Indikatoren auf der Basis theoretischer Überlegungen zur inhaltlichen Validität des Konstruktes beitragen,¹⁷⁷ ist die Güte der Messung insgesamt positiv zu beurteilen.

¹⁷⁷ Vgl. Abschnitt 4.3.1 auf Seite 121.

4.3.1.3 Investierter Anteil der geplanten Ressourcen

Das Konstrukt investierter Anteil der geplanten Ressourcen (Kurzbezeichnung: Ressourceneinsatz) bezieht sich auf die gleichen Ressourcen (Manntage, Kosten, Dauer) wie das vorangegangene Konstrukt. Im Gegensatz zu einer subjektiven Bewertung wird diesmal jedoch um eine Schätzung des investierten Anteils im Vergleich zur Planung gebeten.

Diese Messung folgt der Logik, dass eine stärkere Projekteskalation immer auch einen größeren gesamten Ressourcenverbrauch zur Folge hat. Dies ist eine plausible Annahme, da sich der gesamte Ressourcenverbrauch aus der Summe von (rationalem) Ressourcenverbrauch bis zum rationalen Zeitpunkt des Abbruchs und dem (eskalierten) Ressourcenverbrauch vom rationalen Abbruchzeitpunkt bis zum tatsächlichen Abbruchzeitpunkt ergibt. Von einem negativen Zusammenhang dieser beiden Ressourcenverbräuche ist nicht auszugehen.

Unter diesen Annahmen ist von einer starken Korrelation zwischen Projekteskalation und gesamtem Ressourcenverbrauch auszugehen.

Tabelle 4.7: Gütekriterien des Messmodells „investierter Anteil der geplanten Ressourcen“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Welcher Anteil der ursprünglich (bei Projektbeginn) geplanten Ressourcen wurde bis zum Abbruch des Projekts eingesetzt?				
1.	Manntage ... %	0,81	1,00	fixiert
2.	Kosten (finanzielle Ressourcen) ... %	0,67	0,59	20,36
3.	Dauer des Projekts (Kalenderzeit) ... %	0,62	0,43	17,43
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,82, Varianzerklärung durch den ersten Faktor: 0,76, Faktorreliabilität: 1,00 und DEV: 0,62				

Wird zusätzlich die Annahme getroffen, dass die Budgetplanung zu Beginn des Projekts in etwa den Ressourcen entspricht, die für das Projekt aufzuwenden gewesen wären, wenn keine unerwarteten (und nicht im Puffer berücksichtigten)¹⁷⁸ Probleme aufgetreten wären,¹⁷⁹ so kann davon ausgegangen werden, dass die Projekteskalation noch stärker mit dem Quotient aus gesamtem Ressourcenverbrauch und geplantem Ressourcenverbrauch korreliert. Deshalb wird

¹⁷⁸ Der Effekt der Overconfidence kann die Geltung dieser Annahme einschränken.

¹⁷⁹ Der Effekt des Optimismus kann die Geltung dieser Annahme einschränken.

im Rahmen des Fragebogens der Ressourcen-Einsatz im Vergleich zur Planung abgefragt (in Bezug auf Kosten, Mannstunden und Kalenderzeit).¹⁸⁰

Das Konstrukt zum investierten Anteil der geplanten Ressourcen wurde mit folgendem Hinweis zum Format der Antworten versehen: „Angaben in Prozent der ursprünglich geplanten Gesamtmenge. Also beispielsweise 50 Prozent, wenn bis zum Abbruch die Hälfte der zu Beginn geplanten Ressourcen investiert wurde, oder 200 Prozent, wenn das Doppelte der ursprünglich geplanten Ressourcen investiert wurde.“

Die Gütemaße des Messmodells sind sowohl auf Indikatorebene als auch auf Faktorebene zufriedenstellend. Eine Bereinigung der Items ist nicht notwendig.

4.3.1.4 Diskussion der Projekteskalation als reflektives Konstrukt zweiter Ordnung

Die drei Konstrukte *Verspätung des Projektabbruchs*, *Beurteilung des Ressourcen-Einsatzes* und *investierter Anteil der geplanten Ressourcen* stellen drei Variablen dar, in denen sich die Auswirkung der Projekteskalation widerspiegelt. Dementsprechend kann die Projekteskalation als Konstrukt zweiter Ordnung aufgefasst werden. Darunter werden Konstrukte verstanden, die über andere Konstrukte (erster Ordnung) mit den Indikatoren verknüpft sind.¹⁸¹

Die Verwendung von Konstrukten höherer Ordnung bietet sich an, wenn der gewünschte Grad der Abstraktion einen Sachverhalt betrifft, der sich wiederum in mehreren Facetten widerspiegelt. Gleichzeitig ermöglicht die Verwendung von Konstrukten höherer Ordnung eine Komplexitätsreduktion.¹⁸² Beide Argumente sprechen im Fall der Projekteskalation für eine Modellierung als Konstrukt zweiter Ordnung.

Gewählt wird hier ein *reflektives* Konstrukt zweiter Ordnung.¹⁸³ Für diese Spezifikation spricht, dass die Projekteskalation sich nicht aus den drei Konstrukten erster Ordnung zusammensetzt

¹⁸⁰ Diese Vorgehensweise wenden auch Dilts/Pence an; vgl. Dilts/Pence 2006a, Dilts/Pence 2006b und Dilts/Pence 2006c. Im Unterschied zur vorliegenden Studie nutzen sie diese Angaben im Zusammenhang mit anderen Fragen um einen multidimensionalen Blick auf das mögliche Vorliegen eines Sunk-Cost-Bias zu erhalten; vgl. Dilts/Pence 2006b, S. 386.

¹⁸¹ Vgl. Byrne 2001, S. 37 f.

¹⁸² Vgl. Albers/Götz 2006, S. 672.

¹⁸³ Für eine kritische Einschätzung reflektiver Konstrukte zweiter Ordnung vgl. Albers/Götz 2006, S. 675. Trotz der Kritikpunkte von Albers/Götz sind reflektive Konstrukte zweiter Ordnung gängige Praxis in der internationalen Forschung; vgl. beispielsweise Min/Mentzer/Ladd 2007, S. 510.

(was ein formatives Konstrukt nahelegen würde), sondern stattdessen Auswirkungen auf alle drei Konstrukte erster Ordnung hat. Wenn ein Projekt eskaliert, ist davon auszugehen, dass im Regelfall alle drei Konstrukte gleichzeitig davon beeinflusst werden. Der investierte Anteil der geplanten Ressourcen nimmt zu, die Bewertung der aufgewendeten Ressourcen verschiebt sich in Richtung „zu viel“ und die Beteiligten erleben den Abbruch als zunehmend „zu spät“.

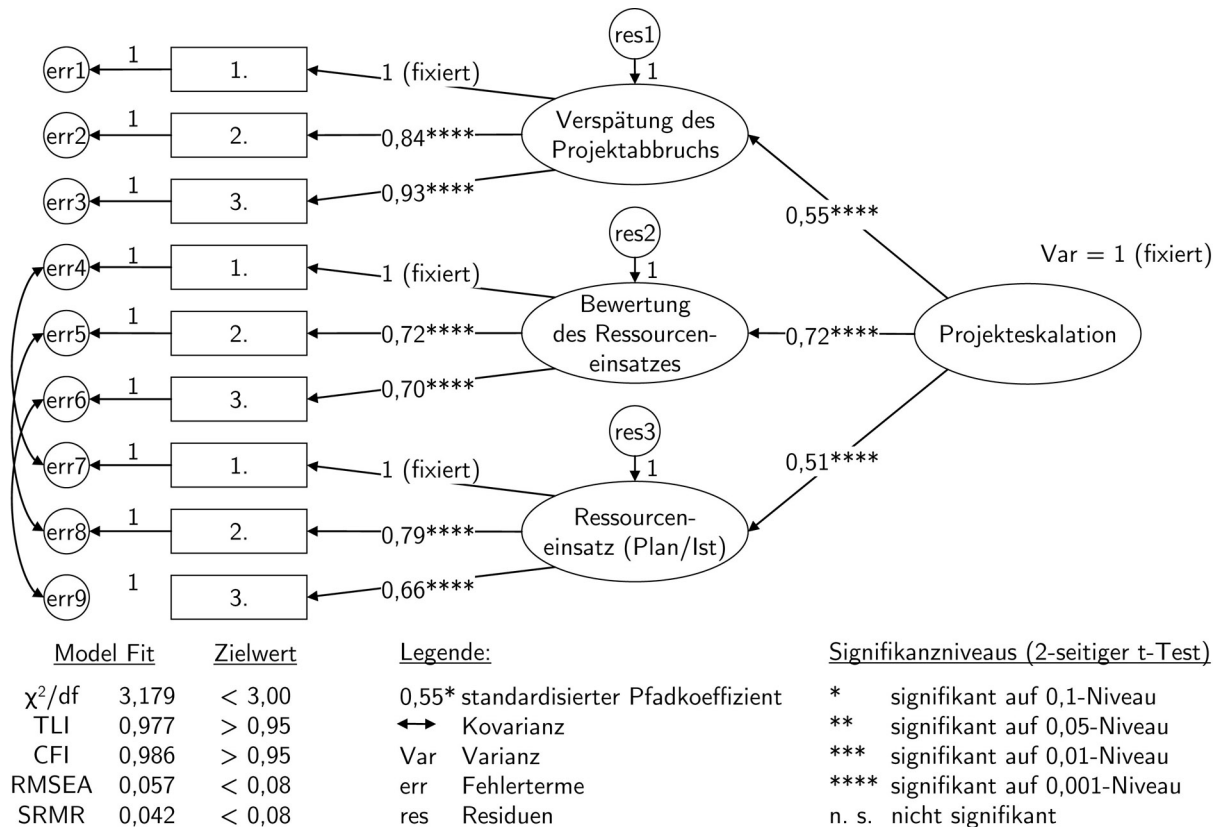


Abbildung 4.3: Projekteskalation als reflektives Konstrukt zweiter Ordnung

Eine konfirmatorische Faktorenanalyse zeigt, dass die Fehlerterme der Indikatoren der Konstrukte *Bewertung des Ressourceneinsatzes* und *investierter Anteil der geplanten Ressourcen*, die sich auf die selbe Ressource beziehen (also Manntage, Kosten oder Zeiteinsatz), signifikante Kovarianzen aufweisen. Aus inhaltlicher Sicht sind die Kovarianzen sehr plausibel, da die Bewertung einer Ressource mit dem Ausmaß des Einsatzes der entsprechenden Ressource korrelieren sollte. Im Messmodell wurden deshalb die Fehlerterme der Konstrukte *Bewertung des Ressourceneinsatzes* und *investierter Anteil der geplanten Ressourcen* jeweils paarweise durch Kovarianzen verbunden, da sich die Fragen stets auf die gleiche Ressource beziehen.¹⁸⁴

¹⁸⁴ Das gleiche Vorgehen wurde bei der konfirmatorischen Faktorenanalyse zur Prüfung der Diskriminanzvalidität gewählt; vgl. Abschnitt 4.3.5.1 auf Seite 144.

Damit ein Modell zweiter Ordnung geschätzt werden kann, muss ebenso wie bei den Messmodellen erster Ordnung entweder einer der Pfadkoeffizienten oder die Varianz des Konstrukts zweiter Ordnung fixiert werden. Im Gegensatz zu Messmodellen erster Ordnung, bei denen üblicherweise *einer der Pfade* auf den Wert eins fixiert wird, ist es bei Modellen zweiter Ordnung in der Regel die *Varianz*, die auf eins fixiert wird.¹⁸⁵ Der Grund dafür liegt darin, dass vor allem die Wirkung des Konstrukts zweiter Ordnung auf die Konstrukte erster Ordnung von Interesse ist. Diese Argumentation trifft auch für den hier vorliegenden Fall zu, weshalb die Varianz fixiert wird. Demzufolge können alle Faktorladungen geschätzt und die Signifikanz ihrer Wirkung ermittelt werden.

Bei Modellen höherer Ordnung muss für jede Ebene einzeln ermittelt werden, ob das Modell genug Freiheitsgrade aufweist.¹⁸⁶ Die Freiheitsgrade eines Modells lassen sich mit der folgenden Formel ermitteln.

$$df = \frac{p \times (p + 1)}{2} - s \quad (4.1)$$

Dabei steht *df* für die Anzahl der Freiheitsgrade, *p* für die Anzahl der beobachteten Variablen und *s* für die Anzahl der zu schätzenden Parameter.¹⁸⁷ Wird die erste Ebene des Messmodells separat (also nur das Konstrukt Projekteskalation und die drei Konstrukte erster Ordnung *ohne* deren Indikatoren) betrachtet, so beträgt die Anzahl der beobachteten Variablen drei und die Anzahl der zu schätzenden Parameter sechs (drei Pfadkoeffizienten plus drei Residuen).

Entsprechend der Formel 4.1 ist das Modell damit gerade identifiziert, besitzt jedoch keine Freiheitsgrade. Das Modell lässt sich damit berechnen, allerdings kann nicht mit Sicherheit gesagt werden, dass der Faktor zweiter Ordnung die Daten besser repräsentiert, als es Korrelationen zwischen den Konstrukten tun würden. Das gesamte Messmodell einschließlich der Indikatoren besitzt 21 Freiheitsgrade.

Zur Prüfung der Güte des Modells wurde der Vorgehensweise von Giere/Wirtz/Schilke gefolgt.¹⁸⁸ Die Gütemaße für das Gesamtmodell zweiter Ordnung (einschließlich der Konstrukte erster Ordnung) sind in Tabelle 4.8 auf der nächsten Seite dargestellt. Das Modell besitzt bei

¹⁸⁵ Vgl. Byrne 2001, S. 38.

¹⁸⁶ Vgl. Byrne 2001, S. 39., und Giere/Wirtz/Schilke 2006, S. 687

¹⁸⁷ Vgl. Byrne 2001, S. 34 ff.

¹⁸⁸ Vgl. Giere/Wirtz/Schilke 2006, S. 686.

Tabelle 4.8: Gütekriterien des Messmodells zweiter Ordnung „Projekteskalation“

Nr.	Faktor erster Ordnung	korrigierte Item-to-Total-Korrelation	t-Wert der Faktorladung
1.	Zu später Abbruch	0,55	9,55
2.	Beurteilung des Ressourcen-Einsatzes	0,72	9,88
3.	Investierte Anteil der geplanten Ressourcen	0,51	9,37

Beurteilung des Gesamtkonstrukts:
 $\frac{\chi^2}{df}$: 3,18, TLI: 0,98, CFI: 0,99, RMSEA: 0,057 und SRMR: 0,042

21 Freiheitsgraden einen χ^2 -Wert von 66,77. $\frac{\chi^2}{df}$ überschreitet damit nur knapp den geforderten Grenzwert von 3,00.¹⁸⁹ Alle anderen Gütemaße sind zufriedenstellend: TLI und CFI liegen über 0,95, RMSEA und SRMR liegen unter 0,08. Eine Modifikation des Messmodells erscheint nicht erforderlich.

4.3.2 Operationalisierung der kognitiven Faktoren

4.3.2.1 Wahrnehmungsschwelle

Damit der Entscheider eine Information wahrnehmen kann, muss diese aus dem „Hintergrundrauschen“ weniger wichtiger Informationen hervorstechen, d. h. eine Schwelle überwinden. Das Konstrukt Wahrnehmungsschwelle soll den Grad der Schwierigkeit für den Entscheidungsverantwortlichen, negative Projektinformationen wahrzunehmen, messen.¹⁹⁰

Die Höhe der Wahrnehmungsschwelle ist dabei von der individuellen Persönlichkeit des Empfängers abhängig. Diese individuelle Abhängigkeit wird in den Indikatoren des Messmodells berücksichtigt, die in Tabelle 4.9 auf der nächsten Seite im Wortlaut aufgeführt werden.

Die Antwortmöglichkeiten liegen, wie auch bei den folgenden Konstrukten, auf einer siebenstufigen Skala von „0 = nie“ bis „6 = immer“.

Für die Fünf-Item-Messung der Wahrnehmungsschwelle sind nicht alle Gütekriterien zufriedenstellend. Der $\frac{\chi^2}{df}$ beträgt 8,38, der RMSEA 0,104. Damit überschreiten die genannten Gütekriterien die üblichen Grenzwerte.

¹⁸⁹ Vgl. Abschnitt 4.2.3 auf Seite 117.

¹⁹⁰ Vgl. Abschnitt 2.3.2.1 auf Seite 30 und die Definition in Tabelle 3.1 auf Seite 54.

Tabelle 4.9: Gütekriterien des Messmodells „Wahrnehmungsschwelle“

Nr.	Wortlaut des Indikators	korrigier- te Item- to-Total- Korrelation	Indikator- reliabilität	t-Wert der Faktorladung
Wichtige negative Informationen ...				
1.	... sind für mich in der Menge der Informationen untergegangen.		eliminiert	
2.	... waren für mich schwer zu bemerken.		eliminiert	
3.	... waren für mich schwer von den unwichtigen Informationen zu unterscheiden.	0,69	0,61	fixiert
4.	... waren zu unauffällig, um von mir gut bemerkt zu werden.	0,68	0,58	18,22
5.	... waren schwer durch mich als solche zu identifizieren.	0,70	0,66	18,54
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,83, Varianzerklärung durch den ersten Faktor: 0,75, Faktorreliabilität: 0,77 und DEV: 0,62				

Ein Grund dafür liegt beim ersten Indikator. Indikator 1 weist eine niedrigere Indikatorreliabilität und eine niedrigere korrigierte Item-to-Total-Korrelation auf als alle anderen Indikatoren. Dies kann damit begründet werden, dass Indikator 1 weniger die Ursache für die Höhe der Wahrnehmungsschwelle abfragt, sondern vielmehr die Folge – nämlich das Vorliegen eines Wahrnehmungsdefizits. Der Indikator wird deshalb aus dem Konstrukt Wahrnehmungsschwelle eliminiert.

Eine andere Ursache für die ungenügenden Gütemaße ist eine hohe Kovarianz zwischen den Fehlertermen von Indikator 2 und Indikator 5, wie eine Analyse der Modifikations-Indizes (Modification Indizes) ergibt. Die Gütekriterien können signifikant verbessert werden, indem entweder eine Kovarianz zwischen beiden Items eingefügt oder eines der beiden Items eliminiert wird. Beide Items liegen inhaltlich nah beieinander, fragen jedoch verschiedene Facetten ab. Bei Indikator 2 geht es darum, ob wichtige negative Informationen *schwer zu bemerken* waren, bei Indikator 5 darum, ob sie *als solche (d. h. als wichtig und negativ) identifiziert* werden konnten. Es ist zu vermuten, dass den Teilnehmern die feine Nuance zwischen beiden Fragen entgangen ist. Basierend auf dieser Annahme, wird eines der beiden Items eliminiert.

Da sich inhaltlich keine Präferenz ergibt, welches der beiden Items eliminiert werden sollte, wurde das Purify Modul des Softwareprogramms Tetrad 4.3.8-0 eingesetzt. Die Tetrad-Analysen ergaben, dass ein Beibehalten von Indikator 5 zu einer leicht höheren durchschnittlich erfassten Varianz führt. Die Wahl des zu eliminierenden Items fiel deshalb auf Indikator 2.

Die Gütekriterien weisen nun auf eine sehr hohe Güte des Messmodells hin, so dass eine weitere Modifikation nicht notwendig ist.

4.3.2.2 Selektive Wahrnehmung

Menschen sind unbewusst bestrebt, ein konsistentes Weltbild aufzubauen. Dementsprechend nehmen sie Informationen bevorzugt wahr, wenn diese mit ihrer bisherigen Einstellung übereinstimmen. Widersprüchliche Argumente werden hingegen weniger in die weitere kognitive Verarbeitung mit einbezogen.¹⁹¹ Mit dem Konstrukt soll gemessen werden, inwiefern relevante Informationen, die den Projekterfolg in Frage gestellt haben, zu spät bemerkt wurden, weil der Entscheider positive Informationen bevorzugt wahrnahm und/oder negative Informationen ausblendete.

Tabelle 4.10: Gütekriterien des Messmodells „selektive Wahrnehmung“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Wenn neue Informationen zum Projekt auftraten, ...				
1.	... schenkte ich positiven Informationen größeres Interesse als negativen.		eliminiert	
2.	... habe ich Informationen weniger Beachtung geschenkt, wenn sie meinen Vorstellungen widersprachen.	0,79	0,73	fixiert
3.	... wandte ich nur bestimmten Informationen meine Aufmerksamkeit zu (d. h., unterlag selektiver Wahrnehmung).	0,79	0,73	25,89
4.	... fokussierte sich meine Aufmerksamkeit auf meine „Lieblingsthemen“.	0,79	0,73	25,91
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,89, Varianzklärung durch den ersten Faktor: 0,82, Faktorreliabilität: 0,85 und DEV: 0,73				

Die Vier-Item-Messung der selektiven Wahrnehmung weist mit einem $\frac{\chi^2}{df}$ von 12,21 einen ungenügenden Fit auf. Die Modification Indices zeigen hohe Kovarianzen von Indikator 1 zu Indikator 2 und 3. Gleichzeitig weist Indikator 1 die geringste Indikatorreliabilität auf. Letzteres könnte daran liegen, dass Indikator 1 nach der selektiven Wahrnehmung *positiver* Fakten fragt, während alle anderen Faktoren selektive Wahrnehmung in Bezug auf eine positive wie auch eine negative Grundhaltung zulassen. Dies spricht dafür, Indikator 1 zu eliminieren. Auch die Analysen mit Tetrad legen nahe, die Items 2 bis 4 zu verwenden. Indikator 1 wird deshalb eliminiert.

Die Gütekriterien zeigen nun eine sehr hohe Güte des Messmodells, so dass eine weitere Modifikation nicht notwendig ist.

¹⁹¹ Vgl. Abschnitt 2.3.2.2 auf Seite 34 und die Definition in Tabelle 3.1 auf Seite 54.

4.3.2.3 Self-Justification-Druck

Das Konstrukt des Self-Justification-Drucks spiegelt die Stärke der empfundenen Bedrohung von Image und Selbstbild bei einem Projektabbruch wider.¹⁹²

Tabelle 4.11: Gütekriterien des Messmodells „Self-Justification-Druck“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Ich hatte Sorge um mein Image, falls es zu einem Projektabbruch kommen würde.	0,81	0,86	fixiert
2.	Ich befürchtete, dass wichtige Personen (z. B. Vorgesetzte, Kapitalgeber u. Ä.) bei einem Projektabbruch einen schlechten Eindruck von mir haben könnten.	0,77	0,73	25,95
3.	Ich stand unter hohem Druck, das Projekt erfolgreich abzuschließen.		eliminiert	
4.	Ich empfand den Erfolg des Projekts als eng verknüpft mit meinem eigenen, persönlichen Erfolg.	0,68	0,52	21,40

Beurteilung des Gesamtkonstrukts:
Cronbachs α : 0,87, Varianzklärung durch den ersten Faktor: 0,80, Faktorreliabilität: 0,89 und DEV: 0,70

Der $\frac{\chi^2}{df}$ -Wert liegt bei der Vier-Item-Messung des Self-Justification-Drucks bei 21,29. Die Ursache lässt sich auf eine hohe Kovarianz zwischen den Fehlertermen von Indikator 3 und 4 zurückverfolgen. Während sich die Indikatoren 1 und 2 auf verschiedene Aspekte der negativen Folgen eines Projektabbruchs fokussieren, bilden Indikator 3 und 4 Aspekte des Bedürfnisses nach Erfolg ab. Während Indikator 3 allgemein nach dem Erfolgsdruck fragt, der also auch extern sein kann, fragt Indikator 4 nach der eigenen persönlichen, internen Sichtweise. möglicherweise wurden die Unterschiede zwischen Indikator 3 und 4 von den Befragten nicht in vollem Umfang wahrgenommen. Für das Self-Justification-Konstrukt weist Indikator 4 eine höhere Bedeutung auf, da er spezifischer formuliert ist als Indikator 3. Gleichzeitig weist Indikator 3 auch die geringste Indikatorreliabilität auf. Sachlogische und statistische Gründe sprechen demzufolge für die Eliminierung von Indikator 3.

Die Gütekriterien zeigen nun eine hohe Güte des Messmodells, so dass eine weitere Modifikation nicht notwendig ist.

¹⁹² Vgl. Abschnitt 2.3.3.1 auf Seite 35 und die Definition in Tabelle 3.1 auf Seite 54.

4.3.2.4 Sunk-Cost-Effekt

Unter Sunk-Cost-Effekt wird die ökonomisch-irrationale Tendenz verstanden, auf Grund bereits erfolgter irreversibler Investitionen Folgeinvestitionen in ein Projekt zu befürworten. Das Konstrukt Sunk-Cost-Effekt misst in der vorliegenden Studie die Empfindung des Entscheidungsträgers, dass die Sunk Costs gegen einen Projektabbruch sprechen.¹⁹³

Tabelle 4.12: Gütekriterien des Messmodells „Sunk-Cost-Effekt“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Einen Projektabbruch sah ich wegen der bereits investierten unwiederbringlichen Investitionen (Sunk Costs) als Verschwendung von Ressourcen an.	0,79	0,79	fixiert
2.	Ich sprach mich für eine Projektfortführung aus, weil bei einem Projektabbruch die zuvor erfolgten Investitionen umsonst gewesen wären.	0,67	0,53	21,51
3.	Einen Projektabbruch habe ich als Vernichtung von bereits investierten Ressourcen betrachtet.	0,77	0,73	26,03
4.	Ein Abbruch des Projekts erschien mir zunehmend teurer, je mehr Ressourcen in das Projekt geflossen waren.	0,60	0,42	18,41

Beurteilung des Gesamtkonstrukts:

Cronbachs α : 0,86, Varianzerklärung durch den ersten Faktor: 0,71, Faktorreliabilität: 0,89 und DEV: 0,62

$\frac{\chi^2}{df}$: 2,64, TLI: 0,99, CFI: 1,00, RMSEA: 0,049 und SRMR: 0,011

Da das Konstrukt mehr als drei Indikatoren umfasst, können auch $\frac{\chi^2}{df}$, TLI, CFI, RMSEA und SRMR für die Beurteilung herangezogen werden. $\frac{\chi^2}{df}$ ist kleiner als 3, TLI und CFI liegen über 0,95 und RMSEA sowie SRMR unterschreiten 0,08.¹⁹⁴ Die Gütekriterien für die Messung der Sunk Costs weisen somit sowohl auf Indikatorebene als auch auf Konstruktebene zufriedenstellende Werte auf. Das Messmodell wurde deshalb nicht modifiziert.

4.3.2.5 Optimismus

Das Konstrukt Optimismus misst den Glauben des Verantwortlichen an den Erfolg des Projekts.¹⁹⁵ Das entsprechende Vier-Item-Konstrukt überschreitet mit einem $\frac{\chi^2}{df}$ -Wert von 21,92 die empfohlene Grenze jedoch bei Weitem. Dabei weist Indikator 1 die geringste

¹⁹³ Vgl. Abschnitt 2.3.4.1 auf Seite 38 und die Definition in Tabelle 3.1 auf Seite 54.

¹⁹⁴ Vgl. Abschnitt 4.2.3 auf Seite 117.

¹⁹⁵ Vgl. Abschnitt 2.3.4.2 auf Seite 40 und die Definition in Tabelle 3.1 auf Seite 54.

Faktorreliabilität auf. Diese liegt mit 0,39 unter dem Grenzwert von 0,4. Eine genaue Betrachtung der Formulierungen ergibt, dass Indikator 1 als einziger explizit nach *übermäßigem* Optimismus fragt: Die Formulierung lautet nämlich „zu positiv“, während alle anderen Indikatoren kein Element wie „zu sehr“ oder „übermäßig“ beinhalten.

Dies spricht dafür, dass Optimismus und übermäßiger Optimismus als zwei unterschiedliche Konstrukte aufgefasst werden müssen. Indikator 1 wird deshalb ausgegliedert und als Single Item für die Variable übermäßiger Optimismus verwendet.¹⁹⁶

Tabelle 4.13: Gütekriterien des Messmodells „Optimismus“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Ich bewertete das Projekt lange Zeit zu positiv.	eliminiert (übermäßiger Optimismus)		
2.	Ich war lange Zeit optimistisch, was den erwarteten Ausgang des Projekts betraf.	0,77	0,68	fixiert
3.	Auch wenn es Probleme gab, war ich zuversichtlich, dass das Projekt am Ende erfolgreich sein würde.	0,82	0,79	26,58
4.	Ich war überzeugt vom erfolgreichen Abschluss des Projekts.	0,81	0,78	26,45

Beurteilung des Gesamtkonstrukts:
Cronbachs α : 0,90, Varianzerklärung durch den ersten Faktor: 0,83, Faktorreliabilität: 0,80 und DEV: 0,75

Die verbleibenden drei Indikatoren zeigen zufriedenstellende Gütemaße. Eine weitere Modifikation ist demnach nicht erforderlich.

4.3.2.6 Übermäßiger Optimismus

Die Variable übermäßiger Optimismus misst den *übermäßigen* Glauben an den Erfolg des Projekts durch den Entscheidungsverantwortlichen.¹⁹⁷ Wie sich im vorherigen Abschnitt gezeigt hat, stellen Optimismus und übermäßiger Optimismus zwei verschiedene Konstrukte dar. Dieses unvorhergesehene Resultat der Validierung der Messmodelle hat Konsequenzen für die Spezifikation des Strukturmodells. Eine erneute Lektüre der Literatur mit dem speziellen Fokus, zwischen Optimismus und übermäßigem Optimismus zu differenzieren, führte zu folgenden Ergebnissen. Zayer verwendet in der Darstellung seines Modells¹⁹⁸ den Begriff „Optimismus“.

¹⁹⁶ Vgl. Abschnitt 4.3.2.6.

¹⁹⁷ Vgl. Abschnitt 2.3.4.2 auf Seite 40.

¹⁹⁸ Vgl. Abbildung 2.2 auf Seite 22.

Ebenso lautet auch die Überschrift des entsprechenden Abschnitts bei Zayer, „Der Effekt des Optimismus“.¹⁹⁹ Im Text des Abschnitts wird gelegentlich von Optimismus gesprochen, jedoch der Begriff *übermäßiger Optimismus* fett hervorgehoben.²⁰⁰ Die referenzierte Literatur verwendet beispielsweise den Begriff „unrealistic optimism“.²⁰¹ Vollzieht man die Argumentation des Effekts nach, der in enger Verbindung mit einer Selbstüberschätzung der Subjekte, dem Hybris-Effekt, gesehen wird,²⁰² so wird deutlich, dass ein Effekt auf die Eskalation durch *übermäßigen Optimismus* zu erwarten ist.

Die im bisherigen Verlauf als Optimismus bezeichnete Variable muss demzufolge bei erneuter Betrachtung als *übermäßiger Optimismus* präzisiert werden. Ebenso muss in den damit verbundenen Hypothesen (H_5 , H_8 , H_{M17} und H_{M18} ; vgl. Abbildung 2.3 auf Seite 28) der Begriff Optimismus um das Adjektiv „übermäßiger“ ergänzt werden. In der vorliegenden Arbeit wird deshalb im Folgenden die Variable *übermäßiger Optimismus* verwendet. Der übermäßige Optimismus wird dabei, wie zuvor erläutert, mit einem Single Item gemessen. Der Wortlaut sowie Mittelwert und Standardabweichung des Items ist in Tabelle 4.14 aufgeführt.

Tabelle 4.14: Variable „übermäßiger Optimismus“

Nr.	Wortlaut des Indikators	M	SD
1.	Ich bewertete das Projekt lange Zeit zu positiv.	3,2	1,6

M = arithmetischer Mittelwert, SD = Standardabweichung

Wie in Abschnitt 3.2.2.5 auf Seite 59 erläutert wurde, sind für latente Konstrukte Messmodelle mit mehreren Indikatoren zu bevorzugen, da diese ermöglichen, den Messfehler zu schätzen. Es stellt sich deshalb die Frage, wie die Validität und Reliabilität der Messung des übermäßigen Optimismus einzuschätzen ist. Generell ist festzustellen, dass Single-Item-Messungen messtheoretisch problematisch sind, die Ergebnisse einer Strukturgleichungsanalyse jedoch nicht per se ungültig werden lassen.²⁰³ Im vorliegenden Fall positiv hervorzuheben ist, dass der Indikator für übermäßigen Optimismus kurz und klar verständlich ist. Dies spricht dafür, dass nur ein verhältnismäßig geringer Messfehler durch unterschiedliche Interpretationen der Aus-

¹⁹⁹ Vgl. Zayer 2007, S. 142.

²⁰⁰ Vgl. Zayer 2007, S. 142.

²⁰¹ Vgl. beispielsweise Weinstein 1980, Weinstein 1987 und Weinstein 1996.

²⁰² Vgl. Zayer 2007, S. 142.

²⁰³ Vgl. Keil et alii 2000, S. 308. Für eine Diskussion der Angemessenheit von Single-Item-Messungen im Bereich Sportmanagement vgl. Kwon/Trail 2005. Für einen Beispiel zur multiplen Regressionsanalyse mit dem Einsatz von Single-Items vgl. beispielsweise Griffith/Harvey 2001, S. 602.

sage zu erwarten ist. Für zukünftige Studien wäre es jedoch empfehlenswert, ein Multi-Item-Konstrukt für den übermäßigen Optimismus zu entwickeln und zu prüfen, ob die Ergebnisse aus der vorliegenden Studie damit repliziert werden können.

4.3.3 Operationalisierung der moderierenden Faktoren

4.3.3.1 Zeitpunkt des ersten Problemauftretens

Unter dem Zeitpunkt der Problemwahrnehmung wird der Zeitpunkt verstanden, an dem erstmals signifikante Probleme auftraten. Dabei sind diejenigen Probleme relevant, die den Projekterfolg in Frage gestellt haben.²⁰⁴

Das Konstrukt umfasst drei Indikatoren. Die Antwortmöglichkeiten liegen auf einer siebenstufigen Skala von „0 = trifft gar nicht zu“ bis „6 = trifft voll zu“. Die Indikatoren sind im Wortlaut in Tabelle 4.15 aufgeführt, ebenso die Gütekriterien.

Tabelle 4.15: Gütekriterien des Messmodells „Zeitpunkt des ersten Problemauftretens“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Die Probleme sind erst lange nach Projektbeginn aufgetreten.	0,77	0,70	fixiert
2.	Die Gefahren für das Projekt kamen erst sehr spät zum Vorschein.	0,76	0,67	24,51
3.	Die ersten auffälligen Schwierigkeiten tauchten erst lange Zeit nach dem Projektstart auf.	0,82	0,83	26,59

Beurteilung des Gesamtkonstrukts:

Cronbachs α : 0,89, Varianzklärung durch den ersten Faktor: 0,82, Faktorreliabilität: 0,88 und DEV: 0,74

Die Gütekriterien des Drei-Item-Konstrukts zum Zeitpunkt des ersten Problemauftretens sind zufriedenstellend. Auf Indikatorebene sind die Reliabilitäten größer als 0,4. Die t-Werte der Faktorladungen sind größer als 2,33. Auf Konstruktebene liegt Cronbachs α über 0,7, die Varianzklärung durch den ersten Faktor über 0,5, die Faktorreliabilität über 0,6 und die DEV über 0,5. Eine Modifikation des Messmodells ist nicht notwendig.

²⁰⁴ Vgl. Abschnitt 2.3.5.1 auf Seite 41 und die Definition in Tabelle 3.1 auf Seite 54.

4.3.3.2 Nicht hinterfragter Freiraum

Das Konstrukt nicht hinterfragter Freiraum misst den Freiraum des Entscheidungsverantwortlichen, seine eigenen Ziele ohne Hinterfragen durch Dritte in die Entscheidung über Abbruch oder Fortsetzung einfließen zu lassen.²⁰⁵

Die Messung umfasst vier Indikatoren, deren Wortlaut in Tabelle 4.16 dargestellt ist. Die Antwortmöglichkeiten liegen, wie auch bei dem folgenden Konstrukt, auf einer siebenstufigen Skala von „0 = nie“ bis „6 = immer“.

Tabelle 4.16: Gütekriterien des Messmodells „nicht hinterfragter Freiraum“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Bei den Entscheidungen über Abbruch oder Fortsetzung des Projekts ...				
1.	... hatte ich einen großen Freiraum, eigene Interessen einfließen zu lassen.	0,62	0,48	fixiert
2.	... musste ich mich nur vor mir selbst rechtfertigen.	0,70	0,63	17,31
3.	... wurden meine Entscheidungen ohne Hinterfragen akzeptiert.	0,61	0,45	15,20
4.	... blieb es mir überlassen, ob ich meine Gründe offenlegen wollte.	0,71	0,66	17,53
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,83, Varianzerklärung durch den ersten Faktor: 0,66, Faktorreliabilität: 0,87 und DEV: 0,57 χ^2_{df} : 0,27, TLI: 1,00, CFI: 1,00, RMSEA: 0,000 und SRMR: 0,004				

Die Gütekriterien des Vier-Item-Konstrukts zum nicht hinterfragten Entscheidungsfreiraum erfüllen alle Anspruchsniveaus. χ^2_{df} ist kleiner als 3, TLI und CFI liegen über 0,95 und RMSEA sowie SRMR unterschreiten 0,08.²⁰⁶ Eine Modifikation des Messmodells ist dementsprechend nicht erforderlich.

4.3.3.3 Overconfidence

Das Konstrukt Overconfidence misst das Ausmaß des Vertrauens in die Zuverlässigkeit der Prognosen des Projekterfolgs.²⁰⁷

²⁰⁵ Vgl. Abschnitt 2.3.5.2 auf Seite 42 und die Definition in Tabelle 3.1 auf Seite 54.

²⁰⁶ Vgl. Abschnitt 4.2.3 auf Seite 117.

²⁰⁷ Vgl. Abschnitt 2.3.5.3 auf Seite 43 und die Definition in Tabelle 3.1 auf Seite 54.

Das Gütemaß $\frac{\chi^2}{df}$ liegt für das Vier-Item-Konstrukt bei 11,40 und überschreitet damit den empfohlenen Richtwert deutlich. Der Fehlerterm des Indikators 4 weist eine verhältnismäßig hohe Kovarianz mit den Fehlertermen von Indikator 1 und 3 auf. Es ist also anzunehmen, dass für die Respondenten die Formulierungen nicht trennscharf waren. Da Indikator 4 eine deutlich geringere Indikatorreliabilität aufweist als die ersten drei Indikatoren, wird er eliminiert.

Tabelle 4.17: Gütekriterien des Messmodells „Overconfidence“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Ich war sehr überzeugt, dass die Schätzungen der zukünftigen Entwicklung korrekt seien.	0,81	0,76	fixiert
2.	Ich war mir sicher in Bezug auf die Prognosen.	0,83	0,81	29,51
3.	Ich war zuversichtlich, dass sich die Schätzungen für den Erfolg des Projekts bestätigen würden.	0,79	0,72	27,63
4.	Ich hatte mich fest auf die Zukunftsprognosen verlassen.		eliminiert	

Beurteilung des Gesamtkonstrukts:
Cronbachs α : 0,91, Varianzklärung durch den ersten Faktor: 0,84, Faktorreliabilität: 0,77 und DEV: 0,76

Die Gütemaße des Drei-Item-Konstrukts sind zufriedenstellend, so dass kein weiterer Modifikationsbedarf vorliegt.

4.3.4 Operationalisierung der Variablen zur Prüfung der Common-Method-Biases

4.3.4.1 Method-Variance-Marker

Zur Messung der Common-Method-Variance wurde der Umfrage ein Method-Variance-Marker beigelegt.²⁰⁸ Dabei handelt es sich um ein Konstrukt, das aus theoretischer Sicht mit mindestens einer Variablen des zu untersuchenden Modells keine Korrelationen aufweist. Dafür wurde das Konstrukt Kapitalmarktorientierung gewählt. Dieses Konstrukt stammt aus einer Untersuchung von Grieshop und wurde dort mit einem Sample von 1067 Controllern validiert.²⁰⁹

Die Indikatoren und Gütemaße sind in Tabelle 4.18 angegeben. Die Antwortmöglichkeiten lagen auf einer siebenstufigen Skala von „0 = trifft gar nicht zu“ bis „6 = trifft voll zu“.

²⁰⁸ Vgl. Abschnitt 3.3.4 auf Seite 66.

²⁰⁹ Vgl. Grieshop 2008.

Tabelle 4.18: Gütekriterien des Messmodells „Kapitalmarktorientierung“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Das Unternehmen war auf den Kapitalmarkt angewiesen.	0,75	0,67	fixiert
2.	Das Unternehmen bemühte sich um eine gute Kommunikation mit Kapitalmarktakteuren (z. B. Aktionären, Analysten, Gläubigern).	0,74	0,64	22,93
3.	Kapitalmarktorientierung war im Unternehmen groß geschrieben.	0,81	0,82	24,58

Beurteilung des Gesamtkonstrukts:
Cronbachs α : 0,88, Varianzerklärung durch den ersten Faktor: 0,81, Faktorreliabilität: 0,89 und DEV: 0,71

Das Drei-Item-Konstrukt Kapitalmarktorientierung erfüllt für alle relevanten Gütemaße die Anforderungen. Es besteht deshalb kein Bedarf einer Modifizierung.²¹⁰

4.3.4.2 Impression-Management-Index

Um die Ausprägung des Impression-Managements zu kontrollieren, wurden im Zuge der Vorstudien²¹¹ fünf Indikatoren aus Paulhus' Impression-Management-Subscale ausgewählt.

Tabelle 4.19: Mittelwerte und Standardabweichungen des Index „Impression-Management“

Nr.	Wortlaut des Indikators	M	SD
1.	Manchmal erzähle ich Lügen, wenn ich muss. (r)	1,9	1,7
2.	Ich vertusche meine Fehler nie.	4,0	1,7
3.	Ich habe schon Schlechtes über einen Freund hinter seinem Rücken gesagt. (r)	1,6	1,7
4.	Ich nehme nie Dinge, die mir nicht gehören.	4,4	2,0
5.	Ich tratsche nicht über anderer Leute Angelegenheiten.	3,7	2,0
Gesamtindex		4,1	1,1

Die Items betreffen nur die Entscheidungsverantwortlichen (N = 287).

M = arithmetischer Mittelwert, SD = Standardabweichung

(r) = Reversed Item; M und SD für das entsprechende Item sind *nicht* reversed angegeben. Für die Berechnung von M und SD für den Gesamtindex wurden die Antworten der entsprechenden Items jedoch reversed.

²¹⁰ Die Analysen, inwieweit sich die Methodenvarianz auf die Ergebnisse der Studie auswirkt, werden in Abschnitt 4.6.2.1 auf Seite 185 beschrieben.

²¹¹ Vgl. Abschnitt 3.4.1.1 auf Seite 76.

Item 1 und 3 sind entgegengesetzt kodiert (reverse coded), da dies auch im englischsprachigen Original von Paulhus der Fall ist.²¹² Die Antwortmöglichkeiten liegen auf einer siebenstufigen Skala von „0 = trifft gar nicht zu“ bis „6 = trifft voll zu“. Gemäß der Vorgehensweise von Paulhus werden die Indikatoren summiert, um die Impression-Management-Tendenz zu messen.²¹³ Gütemaße sind deshalb nicht zu ermitteln.

Die Mittelwerte und Standardabweichungen der Items sind in Tabelle 4.19 aufgeführt.²¹⁴

4.3.4.3 Leniency-Index

Die Messung des Leniency-Bias erfolgt mit einem Index analog zur Messung des Impression-Managements. Tabelle 4.20 gibt den Wortlaut der Indikatoren wieder. Die Indikatoren entsprechen dem Impression-Management-Index mit dem Unterschied, dass sich die Aussagen nicht auf den Respondenten selbst, sondern auf eine andere Person (den Entscheidungsverantwortlichen) beziehen.

Tabelle 4.20: Mittelwerte und Standardabweichungen des Index „Leniency“

Nr.	Wortlaut des Indikators	M	SD
Bitte geben Sie an, inwieweit die folgenden Aussagen auf den Entscheider zutreffen.			
1.	Manchmal erzählt er Lügen, wenn er muss. (r)	2,5	1,8
2.	Er vertuscht seine Fehler nie.	3,0	1,7
3.	Er hat schon Schlechtes über einen Freund hinter dessen Rücken gesagt. (r)	2,0	1,8
4.	Er nimmt nie Dinge, die ihm nicht gehören.	4,0	1,9
5.	Er tratscht nicht über anderer Leute Angelegenheiten.	3,6	1,8
Gesamtindex		3,6	1,2

Die Items betreffen nur die Nicht-Entscheidungsverantwortlichen (N = 390).

M = arithmetischer Mittelwert, SD = Standardabweichung

(r) = Reversed Item; M und SD für das entsprechende Item sind *nicht* reversed angegeben. Für die Berechnung von M und SD für den Gesamtindex wurden die Antworten der entsprechenden Items jedoch reversed.

Da es sich bei der Messung um einen Index handelt, sind keine Gütemaße zu ermitteln. Die Mittelwerte und Standardabweichungen der Items sind in Tabelle 4.20 angegeben.²¹⁵

²¹² Vgl. Paulhus 1984, Paulhus 1986 und Paulhus 1992. Wie in Abschnitt 3.3.4.4 auf Seite 74 diskutiert, wurde für alle *selbst* entwickelten Konstrukte auf entgegengesetzt kodierte Indikatoren verzichtet.

²¹³ Vgl. Bernardi 2006, S. 47, Paulhus 1986 und Paulhus 1992.

²¹⁴ Die Analysen, inwieweit sich das Impression-Management auf die Ergebnisse der Studie auswirkt, werden in Abschnitt 4.6.2.2 auf Seite 187 beschrieben.

²¹⁵ Die Analysen, inwieweit sich der Leniency-Bias auf die Ergebnisse der Studie auswirkt, werden in Abschnitt 4.6.2.3 auf Seite 190 beschrieben.

4.3.4.4 Selektive Wahrnehmung anderer Manager

Das Konstrukt selektive Wahrnehmung anderer Manager dient, im Rahmen des MTMM-Verfahrens, der Prüfung des Social-Desirability-Bias und des Leniency-Bias.²¹⁶ Die Indikatoren sind analog zum Konstrukt selektive Wahrnehmung gewählt, das in Abschnitt 4.3.2.2 auf Seite 132 vorgestellt wurde. Tabelle 4.21 gibt die Indikatoren im Wortlaut wieder.

Tabelle 4.21: Gütekriterien des Messmodells „selektive Wahrnehmung anderer Manager“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
Ein anderer, durchschnittlich guter Manager hätte damals in meiner Situation ...				
1.	... den positiven Informationen mehr Interesse geschenkt als den negativen.		eliminiert	
2.	... Informationen weniger Beachtung geschenkt, wenn sie seinen Vorstellungen widersprochen hätten.	0,76	0,68	fixiert
3.	... nur bestimmten Informationen seine Aufmerksamkeit zugewandt (d. h. wäre selektiver Wahrnehmung unterlegen).	0,80	0,79	25,40
4.	... seine Aufmerksamkeit auf seine „Lieblingsthemen“ fokussiert.	0,78	0,71	24,60
Beurteilung des Gesamtkonstrukts: Cronbachs α : 0,89, Varianzerklärung durch den ersten Faktor: 0,82, Faktorreliabilität: 0,79 und DEV: 0,73				

Die Vier-Item-Messung der selektiven Wahrnehmung anderer Manager weist einen inakzeptabel hohen $\frac{\chi^2}{df}$ -Wert von 19,96 auf. Ebenso wie beim Konstrukt zur selektiven Wahrnehmung des Entscheiders ist vor allem Indikator 1 die Ursache für die mangelnde Güte des Messmodells. Der Fehlerterm von Indikator 1 weist mit den Fehlertermen aller drei anderen Items verhältnismäßig hohe Kovarianzen auf. Gleichzeitig besitzt Indikator 1 die geringste Indikatorreliabilität. Indikator 1 wird deshalb eliminiert. Die modifizierten Messmodelle selektive Wahrnehmung (des Entscheiders) und selektive Wahrnehmung von anderen Managern umfassen die gleichen Indikatoren und bleiben somit vergleichbar.

Die Gütekriterien weisen nun auf eine hohe Güte des Messmodells hin, so dass eine weitere Modifikation nicht notwendig ist. Die Analysen mit dem Konstrukt selektive Wahrnehmung anderer Manager werden in Abschnitt 4.6.2.2 auf Seite 188 diskutiert.

²¹⁶ Vgl. Abschnitt 3.3.4.1 auf Seite 68.

4.3.4.5 Freundschaft

Das Freundschaftskonstrukt wurde erhoben, um den Leniency-Bias zu kontrollieren.²¹⁷ Dabei handelt es sich um mögliche Verzerrungen der Antworten der Nicht-Entscheider, wenn diese den Entscheidungsverantwortlichen in sozial erwünschter Weise beschreiben. Die Stärke der Freundschaft zwischen Respondent und Entscheidungsverantwortlichem ist deshalb von Interesse, weil davon auszugehen ist, dass sie die Stärke des Leniency-Bias mit bedingt.

Da der Leniency-Bias nur bei Auskünften über andere Personen auftritt, wurde das Freundschaftskonstrukt nur bei den Nicht-Entscheidungsverantwortlichen abgefragt. Dementsprechend betreffen die folgenden Analysen nur die 390 Nicht-Entscheidungsverantwortlichen.

Tabelle 4.22: Gütekriterien des Messmodells „Freundschaft“

Nr.	Wortlaut des Indikators	korrigierte Item-to-Total-Korrelation	Indikatorreliabilität	t-Wert der Faktorladung
1.	Ich habe ein gutes persönliches Verhältnis zu dem Entscheider.	0,79	0,90	fixiert
2.	Ich fühle mich mit dem Entscheider freundschaftlich verbunden.	0,62	0,89	13,21
3.	Ich verstehe mich gut mit dem Entscheider.	0,70	0,61	15,29

Beurteilung des Gesamtkonstrukts:
Cronbachs α : 0,84, Varianzklärung durch den ersten Faktor: 0,76, Faktorreliabilität: 0,85 und DEV: 0,80

Die Gütemaße sind zufriedenstellend, eine Modifikation des Konstrukts ist nicht erforderlich. Die Analysen zum Freundschaftskonstrukt werden im Abschnitt zum Leniency-Bias beschrieben (vgl. Abschnitt 4.6.2.3 auf Seite 190).

4.3.5 Diskriminanzvalidität der Konstrukte

Im Anschluss an die Prüfung der Konvergenzvalidität jedes einzelnen Messmodells ist die Diskriminanzvalidität zwischen allen Messmodellen zu prüfen. Dazu können zum einen die Kovarianzen der Residuen im Rahmen einer konfirmatorischen Faktorenanalyse betrachtet werden.²¹⁸ Zum anderen kann das Fornell/Larcker-Kriterium zur Beurteilung herangezogen werden.²¹⁹

²¹⁷ Vgl. Abschnitt 3.3.4.2 auf Seite 72.

²¹⁸ Vgl. Abschnitt 4.3.5.1 auf der nächsten Seite.

²¹⁹ Vgl. Abschnitt 4.3.5.2 auf Seite 146.

4.3.5.1 Konfirmatorische Faktorenanalyse

Die Prüfung der Diskriminanz kann, ebenso wie die Konvergenzvalidität, mit der konfirmatorischen Faktorenanalyse erfolgen.²²⁰ Um die Diskriminanzvalidität zu prüfen, wurden alle Konstrukte simultan geschätzt, wobei Kovarianzen zwischen allen Konstrukten zugelassen wurden.²²¹ Anschließend wurden die standardisierten residualen Kovarianzen betrachtet. Die residualen Kovarianzen entsprechen der Differenz aus der restringierten Kovarianz-Matrix des hypothetisierten Modells ($\sum(\theta)$) und der Sample-Kovarianz-Matrix (S). Die Matrix der residualen Kovarianzen ($\sum(\theta) - S$) gibt für jede Kombination zweier Indikatoren an, wie stark das Modell von den empirischen Daten abweicht. Werte von über 2,58 gelten dabei als signifikant.²²² Von allen Items weisen nur die Indikatoren 1 und 3 des Konstrukts investierter Anteil der geplanten Ressourcen signifikante standardisierte residuale Kovarianzen auf.

Zusätzlich wurden die von Amos ausgegebenen Modification Indizes betrachtet.²²³ Diese geben an, wie stark sich der Fit des Modells verbessert, wenn zusätzliche Regressionspfade oder Kovarianzen berücksichtigt werden. Modifikationen sollten jedoch nur vorgenommen werden, wenn sie auch substanztheoretisch sinnvoll sind. Im vorliegenden Fall legen die Modifikationsindizes nahe, den Fehlerterm von Indikator 2 von „Beurteilung des Ressourcen-Einsatzes“ mit dem Fehlerterm von Indikator 2 von „Anteil der geplanten Ressourcen“ kovariieren zu lassen. Ebenso signalisieren die Modifikationsindizes, eine Kovarianz zwischen den Fehlertermen der beiden Indikatoren 3 der beiden genannten Konstrukte zuzulassen. Aus inhaltlicher Sicht sind die Kovarianzen sehr plausibel, da beide Konstrukte sich auf die gleichen Ressourcen beziehen. Deshalb werden die Fehlerterme aller drei Items mit Kovarianzen verbunden, wie dies auch in Abschnitt 4.3.1.4 auf Seite 127 beschrieben wurde.

Tabelle 4.23 auf der nächsten Seite zeigt, welche Korrelationen zwischen den Konstrukten signifikant sind. Die Gütekriterien für die simultane Schätzung aller Messmodelle weisen auf einen sehr guten Fit hin (χ^2_{df} : 0,82, TLI: 0,96, CFI: 0,96, RMSEA: 0,035 und SRMR: 0,034).

²²⁰ Vgl. Hildebrandt/Temme 2006, S. 621 f.

²²¹ Vgl. Byrne 2001, S. 88 ff. Dabei wurde das Messmodell zweiter Ordnung zur Kontrolle zunächst ohne Kovarianzen zwischen den Indikatoren spezifiziert (vgl. Abschnitt 4.3.1.4 auf Seite 127).

²²² Vgl. Byrne 2001, S. 89.

²²³ Vgl. Byrne 2001, S. 90 f.

Tabelle 4.23: Unterscheidung zwischen signifikanten und nicht signifikanten Korrelationen zwischen den Konstrukten

Konstrukt	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Verspätung des Projektabbruchs	1,00												
2. Beurteilung des Ressourcen-Einsatzes	0,40	1,00											
3. Investierter Anteil der geplanten Ressourcen	0,28	0,37	1,00										
4. Wahrnehmungsschwelle	0,29	0,15	0,10	1,00									
5. selektive Wahrnehmung	0,22	-0,05	0,04	0,30	1,00								
6. Self-Justification-Druck	0,28	0,11	0,10	0,25	0,45	1,00							
7. Sunk-Cost-Effekt	0,41	0,18	0,18	0,33	0,26	0,50	1,00						
8. Optimismus	0,14	0,04	0,08	0,15	0,17	0,28	0,35	1,00					
9. Übermäßiger Optimismus	0,35	0,13	0,08	0,28	0,42	0,38	0,35	0,61	1,00				
10. Zeitpunkt des ersten Problemauftretens	0,10	0,15	0,15	0,30	-0,06	0,02	0,19	0,28	0,13	1,00			
11. Nicht hinterfragter Freiraum	-0,03	-0,07	-0,05	0,00	0,29	-0,03	0,10	0,15	0,09	0,09	1,00		
12. Overconfidence	0,00	-0,07	0,00	0,09	0,17	0,07	0,19	0,57	0,35	0,21	0,29	1,00	
13. Kapitalmarktorientierung	0,08	0,01	-0,02	0,01	0,10	0,19	0,03	0,07	0,00	0,03	-0,08	0,05	1,00
14. selektive Wahrnehmung anderer Manager	0,17	-0,01	-0,06	0,40	0,41	0,23	0,19	0,07	0,14	0,08	0,05	0,01	0,07

Alle Korrelationen sind signifikant ($p < 0,05$), abgesehen von den *kursiv* gestellten.

Der Impression-Management-Index wird hier nicht aufgeführt, da es sich um einen Index und nicht um ein latentes Konstrukt handelt.

Das Freundchaftskonstrukt wird hier nicht aufgeführt, da es nur die Hälfte des Datensatzes (Nicht-Entscheidungsverantwortliche) betrifft und somit nicht direkt mit den anderen Konstrukten vergleichbar ist.

4.3.5.2 Fornell/Larcker-Kriterium

Das Fornell/Larcker-Kriterium stellt ein Maß zur Bewertung der Diskriminanzvalidität dar. Es besagt, dass Diskriminanzvalidität vorliegt, wenn die quadrierten Korrelationen zwischen zwei Konstrukten kleiner sind als die durchschnittlich erfassten Varianzen jedes der beiden Konstrukte.²²⁴ Formal dargestellt lautet das Fornell/Larcker-Kriterium wie folgt:

$$\text{DEV}(\xi_i) > \text{quadrierte Korrelation}(\xi_i, \xi_j), \text{ für alle } i \neq j \quad (4.2)$$

Das Fornell/Larcker-Kriterium ist geeignet, um Multikollinearität nachzuweisen.²²⁵ Homburg/Klarmann weisen darauf hin, dass das Fornell/Larcker-Kriterium konzeptionelle Überlegungen bei der Auswahl der Konstrukte nicht ersetzen kann, sondern ergänzen sollte.²²⁶ Der Aspekt der Diskriminanzvalidität wurde deshalb in der vorliegenden Arbeit schon in der frühen Phase der Modifikation des Modells adressiert.²²⁷

Die Ergebnisse der Prüfung des Fornell/Larcker-Kriteriums sind in Tabelle 4.24 auf der nächsten Seite dargestellt. Die quadrierte Korrelation zwischen je zwei Faktoren ist in allen Fällen kleiner als die jeweiligen durchschnittlich erfassten Varianzen der beiden entsprechenden Konstrukte. Damit ist die Diskriminanzvalidität der Konstrukte gemäß dem Fornell/Larcker-Kriterium erfüllt.

²²⁴ Vgl. Fornell/Larcker 1981, S. 39 ff.

²²⁵ Vgl. Grewal/Cote/Baumgartner 2004, S. 528.

²²⁶ Vgl. Homburg/Klarmann 2006, S. 729 f.

²²⁷ Vgl. hierzu die Ausführungen auf Seite 27.

Tabelle 4.24: Diskriminanzvalidität anhand des Fornell/Larcker-Kriteriums

Konstrukt	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
	DEV	0,80	0,53	0,62	0,62	0,73	0,70	0,62	0,75	0,74	0,57	0,76	0,71
1. Verspätung des Projektabbruchs	0,80												
2. Beurteilung des Ressourcen-Einsatzes	0,53	0,16											
3. Investierter Anteil der geplanten Ressourcen	0,62	0,08	0,13										
4. Wahrnehmungsschwelle	0,62	0,09	0,02	0,01									
5. selektive Wahrnehmung	0,73	0,05	0,00	0,00	0,09								
6. Self-Justification-Druck	0,70	0,08	0,01	0,01	0,06	0,21							
7. Sunk-Cost-Effekt	0,62	0,16	0,03	0,03	0,11	0,07	0,25						
8. Optimismus	0,75	0,02	0,00	0,01	0,02	0,03	0,08	0,13					
9. Übermäßiger Optimismus	1*	0,12	0,02	0,01	0,08	0,18	0,14	0,12	0,38				
10. Zeitpunkt des ersten Problemauftretens	0,74	0,01	0,02	0,02	0,09	0,00	0,04	0,08	0,02	0,02			
11. Nicht hinterfragter Freiraum	0,57	0,02	0,00	0,00	0,00	0,08	0,01	0,02	0,02	0,01	0,08		
12. Overconfidence	0,76	0,00	0,00	0,00	0,01	0,03	0,00	0,04	0,32	0,12	0,04	0,08	
13. Kapitalmarktorientierung	0,71	0,01	0,00	0,00	0,00	0,01	0,03	0,00	0,00	0,00	0,01	0,00	0,00
14. selektive Wahrnehmung anderer Manager	0,73	0,03	0,00	0,00	0,16	0,17	0,05	0,04	0,00	0,02	0,01	0,00	0,01

DEV = Durchschnittlich erklärte Varianz.

* Fixiert auf 1, da die Messung nur mit einem Item erfolgt.

Da die Korrelationen in Tabelle 4.23 gerundet wurden, weichen die hier angegebenen Werte vereinzelt vom Quadrat der Werte in Tabelle 4.23 ab.

Der Impression-Management-Index wird hier nicht aufgeführt, da es sich um einen Index und nicht um ein latentes Konstrukt handelt.

Das Freundschaftskonstrukt wird hier nicht aufgeführt, da es nur die Hälfte des Datensatzes (Nicht-Entscheidungsverantwortliche) betrifft und somit nicht direkt mit den anderen Konstrukten vergleichbar ist.

4.4 Analyse von Strukturmodellen der Projekteskalation

4.4.1 Analyse der kognitiven Faktoren

Nachdem die Messmodelle in den vorherigen Abschnitten geprüft und, wenn nötig, bereinigt wurden, können die Strukturmodelle untersucht werden. Dabei geben Backhaus/Blechsmidt/Eisenbeiß 2006 zu bedenken: „Die Möglichkeit der Analyse komplexer Kausalstrukturen bringt allerdings das schwerwiegende Problem mit sich, daß [sic] die Zahl der möglichen Modellstrukturen sehr groß wird.“²²⁸ Bei der Analyse von Strukturgleichungsmodellen muss deshalb beachtet werden, dass ein guter Fit mit den Daten nicht bedeutet, dass „das richtige Modell“ gefunden wurde. Häufig gibt es mehrere Modelle, die einen guten Fit mit den Daten aufweisen.²²⁹ Äquivalente Modelle und nichtäquivalente Alternativmodelle nicht zu berücksichtigen, wird als ein typischer Fehler in der Anwendung von SEM angesehen.²³⁰ Im Folgenden werden deshalb verschiedene plausible Modelle untersucht. Zur Komplexitätsreduzierung wird an dieser Stelle zunächst auf die moderierenden Effekte verzichtet; diese werden in Abschnitt 4.4.2 auf Seite 162 separat analysiert.

4.4.1.1 Eskalationsmodell ohne Kovarianzen zwischen den Faktoren (Strukturmodell 1)

Wie in Abschnitt 2.2.2.2 auf Seite 25 ausgeführt, ist bei der Betrachtung realer Projekte davon auszugehen, dass zwischen den verschiedenen kognitiven Faktoren Abhängigkeiten auftreten können. Außerdem ist es möglich, dass mehrere kognitive Faktoren gemeinsam Einflüssen außerhalb des Modells unterliegen. Bei der Datenanalyse ist deshalb von Kovarianzen zwischen den Faktoren auszugehen. Bevor ein solches Modell in Abschnitt 4.4.1.2 auf Seite 150 untersucht wird, soll jedoch zunächst die Gegenhypothese geprüft werden, dass alle Faktoren voneinander unabhängig sind. Dazu wird das Modell in Abbildung 4.4 auf der nächsten Seite geschätzt. In diesem Modell sind keine Kovarianzen modelliert. Der Datensatz wird somit unter der Restriktion geschätzt, dass weder kausale Wirkungen zwischen den kognitiven Faktoren

²²⁸ Backhaus/Blechsmidt/Eisenbeiß 2006, S. 712.

²²⁹ Vgl. Backhaus/Blechsmidt/Eisenbeiß 2006, S. 712.

²³⁰ Vgl. Kline 2005, S. 323.

bestehen noch Faktoren außerhalb des Modells existieren, die mehrere kognitive Faktoren im Modell beeinflussen.

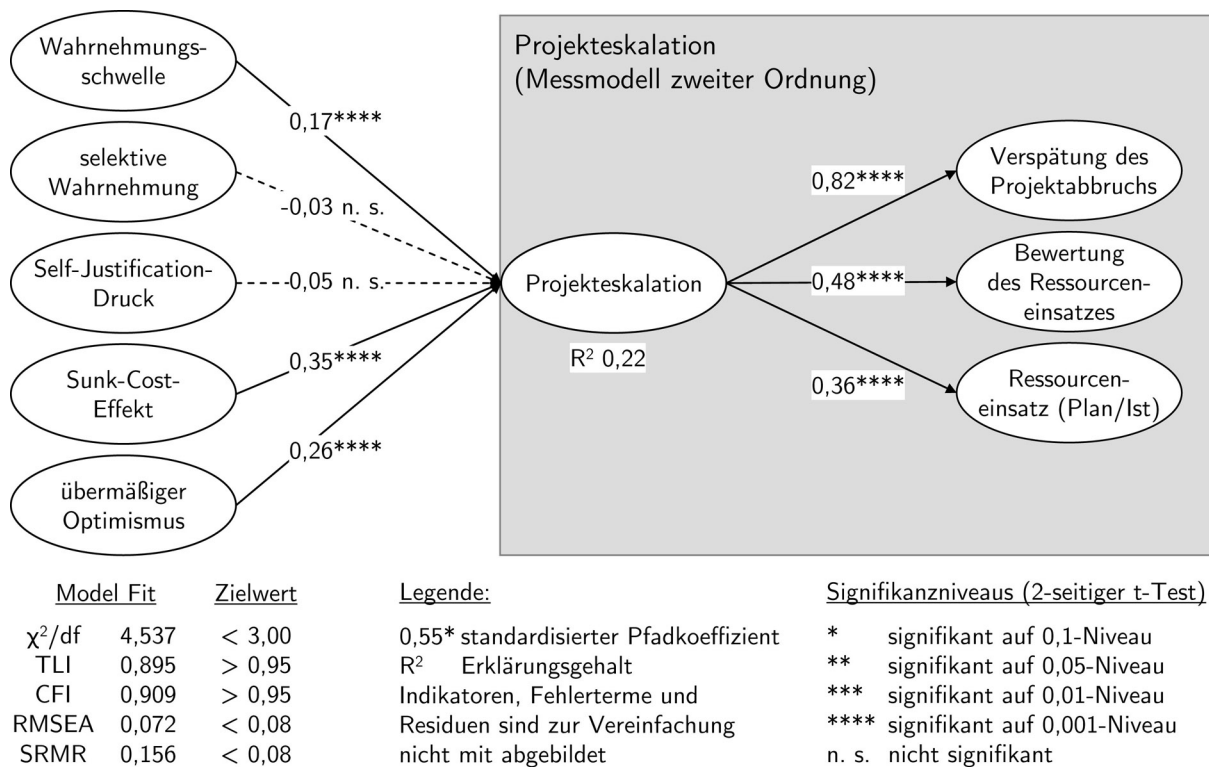


Abbildung 4.4: Modell ohne Kovarianzen zwischen den kognitiven Faktoren (Modell 1)

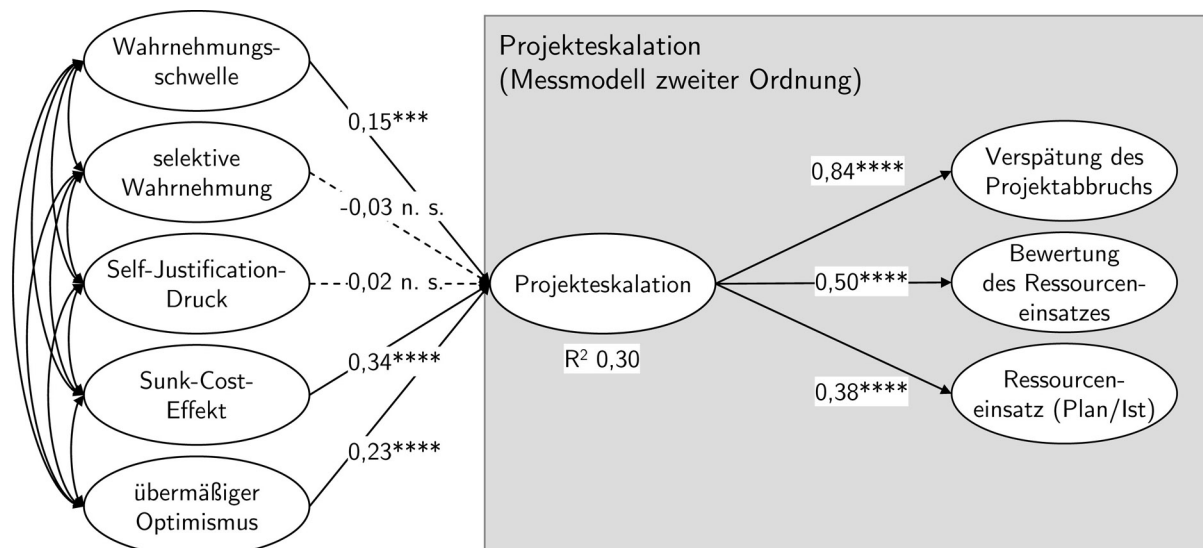
Angeichts der starken Restriktionen ist es nicht überraschend, dass fast alle Gütekriterien einen unzureichenden Fit zwischen empirischen Daten und theoretischem Modell aufzeigen.²³¹ Das Modell weist 220 Freiheitsgrade und einen χ^2 -Wert von 998,05 auf. $\frac{\chi^2}{df}$ beträgt somit 4,537 und überschreitet den Grenzwert von 3,00.²³² TLI und CFI liegen unter 0,95, der SRMR überschreitet 0,08. Lediglich der RMSEA erfüllt mit 0,072 den geforderten Richtwert von 0,08. Insgesamt sprechen die Gütemaße dafür, die Gegenhypothese einer vollständigen Unabhängigkeit zwischen den Faktoren abzulehnen. Da die Gütekriterien keinen zufriedenstellenden Fit belegen, wird auf eine Interpretation der Pfadkoeffizienten verzichtet.

²³¹ Vgl. Abbildung 4.4. Für die Diskussion der Gütemaße vgl. Abschnitt 4.2.3 auf Seite 117.

²³² Vgl. Abschnitt 4.2.3 auf Seite 117.

4.4.1.2 Eskalationsmodell mit Kovarianzen zwischen den Faktoren (Strukturmodell 2)

Das zweite hier untersuchte Modell ist dem ersten Modell sehr ähnlich. Als einziger Unterschied wurde die strenge Annahme vollkommener Unabhängigkeit der kognitiven Faktoren aufgehoben. Stattdessen bildet das Modell die Annahme ab, dass in realen Projekten, die eine Vielzahl einzelner Entscheidungsprozesse beinhalten können, zwischen allen kognitiven Faktoren auch Kovarianzen auftreten können.²³³ Diese Kovarianzen können durch Kausalbeziehungen (einseitig oder wechselseitig) zwischen den kognitiven Effekten hervorgerufen werden oder durch nicht im Modell berücksichtigte Faktoren, die mehrere Modellvariablen gleichzeitig beeinflussen. In Abbildung 4.5 werden die potentiellen Kovarianzen zwischen den kognitiven Faktoren als zweiseitige, geschwungene Pfeile dargestellt.



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df	2,068	< 3,00	0,55* standardisierter Pfadkoeffizient
TLI	0,968	> 0,95	R ² Erklärungsgehalt
CFI	0,974	> 0,95	Indikatoren, Fehlerterme und
RMSEA	0,040	< 0,08	Residuen sind zur Vereinfachung
SRMR	0,046	< 0,08	nicht mit abgebildet
			* signifikant auf 0,1-Niveau
			** signifikant auf 0,05-Niveau
			*** signifikant auf 0,01-Niveau
			**** signifikant auf 0,001-Niveau
			n. s. nicht signifikant

Abbildung 4.5: Modell mit Kovarianzen zwischen den kognitiven Faktoren (Modell 2)

Das Modell 2 in Abbildung 4.5 weist einen guten Fit mit den empirischen Daten auf. Alle Kovarianzen sind auf einem Niveau von 0,001 (beidseitiger Test) signifikant. Ein Blick auf die Gütemaße zeigt, dass alle Werte deutlich besser als die vorgegebenen Schwellenwerte sind.

²³³ Vgl. Abschnitt 2.2.2.2 auf Seite 25.

Das Modell weist 210 Freiheitsgrade und einen χ^2 -Wert von 434,20 auf. $\frac{\chi^2}{df}$ beträgt somit 2,068 und erfüllt den Grenzwert von 3,00.²³⁴ TLI und CFI liegen über 0,95, RMSEA und SRMR unterschreiten 0,08.

Die standardisierten Pfadkoeffizienten sind für die Wirkungen von Wahrnehmungsschwelle, Sunk-Cost-Effekt und übermäßigem Optimismus stark signifikant. Die Hypothesen H₁, H₄ und H₅ werden somit gestützt.

Die selektive Wahrnehmung wirkt nicht signifikant auf die Projekteskalation. Auch der Self-Justification-Druck wirkt im Gesamtsample zunächst nicht signifikant auf die Projekteskalation. (Dies ändert sich jedoch, sobald der von Zayer konstatierte moderierende Effekt des Frei-raums mit einbezogen wird. Dies wird in Abschnitt 4.4.2.2 auf Seite 164 näher betrachtet.) Die Hypothesen H₂ und H₃ müssen somit verworfen werden. Daraus sollte jedoch nicht mit Sicherheit geschlossen werden, dass diese Effekte bei der Projekteskalation keine Rolle spielen.²³⁵

In der Literatur wird beispielsweise darauf hingewiesen, dass Konstrukte der Eskalationstheorie schwer zu operationalisieren und zu messen sind.²³⁶ „Thus, it may be that these constructs *are* equally important, but that is because of the difficulty of measuring them they are less predictive of escalation than one might think.“²³⁷ Dies könnte auch im vorliegenden Fall ein Grund für den nicht signifikanten Zusammenhang sein. Da das Messmodell der selektiven Wahrnehmung jedoch einen guten Fit aufwies, können die Resultate dennoch als starkes Indiz dafür interpretiert werden, dass die selektive Wahrnehmung und der Self-Justification-Druck im Vergleich zu den anderen kognitiven Faktoren eine geringere Rolle spielen, was die *direkte* Wirkung auf die Eskalation betrifft (für mögliche *indirekte* Wirkungen vgl. die Abschnitte 4.4.1.3 und 4.4.1.4).

Der R²-Wert in Abbildung 4.5 zeigt, dass 30 Prozent der Varianz der Projekteskalation durch das Modell der *kognitiven Faktoren* erklärt werden können.²³⁸ Dies ist ein verhältnismäßig ho-

²³⁴ Vgl. Abschnitt 4.2.3 auf Seite 117.

²³⁵ An dieser Stelle sei auf die „Duhem-Quine underdetermination thesis“ hingewiesen, die für die gesamte Interpretation empirischer Daten berücksichtigt werden muss. Diese These besagt, dass Theorien niemals isoliert getestet werden können. Um einen empirischen Test durchzuführen, muss stets auf eine Reihe von zusätzlichen Nebenhypothesen zurückgegriffen werden, beispielsweise über die empirischen Daten, die Testmethode, den Geltungsbereich etc. Dies führt dazu, dass im Fall widersprüchlicher Resultate nicht mit Sicherheit entschieden werden kann, ob das Problem in der ursprünglichen Theorie oder in einer (oder mehreren) der zusätzlichen Nebenhypothesen liegt; vgl. Hands 2001, S. 96.

²³⁶ Vgl. Keil et alii 2003, S. 260.

²³⁷ Keil et alii 2003, S. 260; Hervorhebung im Original.

²³⁸ R² entspricht dem Anteil der Varianz der abhängigen Variablen, der auf die Varianz der unabhängigen Variablen zurückgeführt werden kann.

her Anteil, wenn berücksichtigt wird, dass auch andere Faktoren wie organisationale Verankerung, Projektspezifika, weitere Entscheidungsbefugte und soziale Kontextfaktoren als starke Determinanten der Projekteskalation betrachtet werden können.²³⁹

Zusammenfassend lässt sich an dieser Stelle Folgendes festhalten: Unter Verwendung des Modells von Zayer können (durch Modellierung von Kovarianzen zwischen den Faktoren) die zuvor experimentell und sachlogisch begründeten direkten Wirkungen von drei der fünf untersuchten kognitiven Faktoren auf die Projekteskalation im realen Unternehmenskontext empirisch nachgewiesen werden. Die beiden stärksten Effekte gehen dabei in der vorliegenden Untersuchung vom Sunk-Cost-Effekt und dem übermäßigen Optimismus aus. Forschungsfrage 1 kann damit als beantwortet gelten.²⁴⁰

4.4.1.3 Eskalationsmodell unter Berücksichtigung der kognitiven Dissonanz (Strukturmodell 3)

Mit der Analyse des Modells²⁴¹ und der Betrachtung der moderierenden Effekte²⁴² könnte die Prüfung des Modells von Zayer abgeschlossen werden. Da jedoch, wie zuvor angesprochen,²⁴³ Daten häufig durch verschiedene Modelle statistisch gleichermaßen gut erklärt werden können, sollen an dieser Stelle noch weitere Modelle analysiert werden.²⁴⁴ Dies erscheint auch deshalb geboten, um im Forschungsprozess nicht selbst den besprochenen Effekten der selektiven Wahrnehmung und dem Confirmation Bias zu erliegen.²⁴⁵

Um ein alternatives Modell theoretisch fundiert zu entwickeln, wurde bei den Ursprüngen der Self-Justification-Theorie angesetzt, die ein zentrales Element der Eskalationsforschung darstellt.²⁴⁶ Die Self-Justification-Theorie gründet sich auf die Theorie der kognitiven Dissonanz,²⁴⁷ die wiederum auf Studien von Festinger/Riecken/Schachter zurückgeht.²⁴⁸ Festinger

²³⁹ Vgl. Keil 1995, S. 436, Staw/Ross 1987a, S. 39 ff., und die Erläuterungen in Abschnitt 2.1.2 auf Seite 12.

²⁴⁰ Vgl. für die Forschungsfragen Abschnitt 1.3 auf Seite 7.

²⁴¹ Vgl. Abbildung 4.5.

²⁴² Vgl. Abschnitt 4.4.2 auf Seite 162.

²⁴³ Vgl. Abschnitt 4.4.1 auf Seite 148.

²⁴⁴ Wouters/Anderson/Wynstra merken dazu an: "Models are never confirmed by data; rather, they gain support by failing to be disconfirmed"; Wouters/Anderson/Wynstra 2005, S. 187.

²⁴⁵ Vgl. hierzu Abschnitt 2.3.2.2 auf Seite 34.

²⁴⁶ Vgl. Abschnitt 2.3.3.1 auf Seite 35.

²⁴⁷ Vgl. Keil et alii 1995, S. 68.

²⁴⁸ Vgl. Festinger/Riecken/Schachter 1956.

postuliert zwei Basishypothesen:²⁴⁹ (1) Die Existenz einer Dissonanz wird als unangenehm empfunden und motiviert die Person, die Dissonanz zu verringern und Konsonanz herzustellen. (2) Während die Dissonanz existiert, wird die Person zusätzlich zu dem Versuch des Reduzierens alle Situationen und Informationen aktiv vermeiden, die zu einer Verstärkung der Dissonanz führen.

Nach Festinger existieren drei Möglichkeiten, um die kognitive Dissonanz zu reduzieren: die Veränderung des eigenen Verhaltens, die Veränderung der Umwelt und das Hinzufügen neuer kognitiver Elemente, d. h. eine Veränderung der eigenen Einstellung.²⁵⁰ Die drittgenannte Möglichkeit wird im Rahmen der Forced-Compliance-Forschung untersucht, die zu folgendem Schluss kommt: Wenn Menschen gezwungen werden, eine Position zu vertreten, die ihren eigenen Überzeugungen widerspricht (profan ausgedrückt „lügen müssen“), führt dies dazu, dass sich die eigene Position bzw. Einstellung der geäußerten annähert, denn dadurch kann die kognitive Dissonanz reduziert werden. Dieser Effekt tritt umso stärker auf, je weniger rationale Gründe (beispielsweise monetäre Entlohnung) für das Lügen bestanden haben.²⁵¹

Für die Entwicklung eines Alternativmodells können aus diesen Erkenntnissen folgende Hypothesen abgeleitet werden. Ein hoher Self-Justification-Druck geht mit dem Auftreten einer kognitiven Dissonanz einher. Diese besteht darin, dass der Verantwortliche einerseits Druck verspürt, das Projekt weiterzuführen, um sein Image in den Augen anderer und sein Bild von sich selbst zu schützen. Andererseits spricht das negative Projektfeedback aus Unternehmenssicht dafür, das Projekt zu beenden. Entsprechend der Theorie der kognitiven Dissonanz ist der Entscheidungsträger bestrebt, diesen Konflikt zu verringern.

Dazu stehen ihm die verschiedenen genannten Möglichkeiten zur Verfügung. Eine besteht darin, die positive Einschätzung über die Erfolgsaussichten des Projekts zu verstärken. Dies kann beispielsweise dadurch geschehen, dass negative Informationen aus dem Weg gegangen wird. So stellt Festinger fest, dass Individuen die Wahrnehmung von Informationen vermeiden, die zu einer Vergrößerung der kognitiven Dissonanz führen können.²⁵² Dementsprechend erscheint es plausibel anzunehmen, dass ein höherer Self-Justification-Druck zu einer erhöhten Wahrnehmungsschwelle und zu stärkerer selektiver Wahrnehmung führt.

H_{D1}: Der Self-Justification-Druck wirkt direkt positiv auf die Wahrnehmungsschwelle.

²⁴⁹ Vgl. Festinger 1957, S. 3.

²⁵⁰ Vgl. Festinger 1957, S. 18 ff.

²⁵¹ Vgl. Festinger/Carlsmith 1959, S. 207.

²⁵² Vgl. Festinger 1957, S. 18 ff.

H_{D2}: Der Self-Justification-Druck wirkt direkt positiv auf die selektiver Wahrnehmung.

Neben der Wahrnehmung kann der Entscheidungsverantwortliche (unbewusst) auch seine Einschätzung der Projektbewertung ändern. Beispielsweise könnte er die Sunk Costs als Argument für eine Weiterführung heranziehen oder seine Einschätzung der Erfolgsaussichten, d. h. seinen Optimismus, verstärken. Forschungsergebnisse haben gezeigt, dass Menschen überdies dazu neigen, dissonante Informationen zu vergessen.²⁵³ Dies führt dazu, dass sich die Tendenz zum übermäßigen Optimismus weiter vergrößert.²⁵⁴

Es erscheint demzufolge plausibel, dass ein höherer Self-Justification-Druck zu einem höheren Sunk-Cost-Effekt und stärkerem übermäßigem Optimismus führt.

H_{D3}: Der Self-Justification-Druck wirkt direkt positiv auf den Sunk-Cost-Effekt.

H_{D4}: Der Self-Justification-Druck wirkt direkt positiv auf den übermäßigen Optimismus.

Die Hypothesen H_{D1} bis H_{D4} werden im Modell in Abbildung 4.6 auf der nächsten Seite durch entsprechende Pfeile dargestellt. Für das abgebildete Modell wurden Kovarianzen zwischen den kognitiven Effekten spezifiziert, die durch geschwungene Doppelpfeile repräsentiert werden.

Auch das aus der Theorie der kognitiven Dissonanz begründete Modell 3 zeigt einen guten Fit mit den Daten.²⁵⁵ Das Modell weist 210 Freiheitsgrade und einen χ^2 -Wert von 442,10 auf. $\frac{\chi^2}{df}$ beträgt somit 2,105 und erfüllt den Grenzwert von 3,00.²⁵⁶ TLI und CFI liegen über 0,95, RMSEA und SRMR unterschreiten 0,08.

Aus den Analysen folgt, dass auch das Modell 3 auf der Basis der kognitiven Dissonanz mit den empirischen Daten im Einklang steht und eine zulässige Interpretationsmöglichkeit darstellt. Der Self-Justification-Druck wirkt auf alle anderen kognitiven Faktoren stark signifikant mit standardisierten Pfadkoeffizienten von 0,27 bis 0,50. Die Wirkungen aller Faktoren auf die Projekteskalation entsprechen im Wesentlichen den Ergebnissen aus Modell 2. Auch die erklärte Varianz der Projekteskalation ist mit 0,30 in Modell 3 genauso hoch wie in Modell 2.

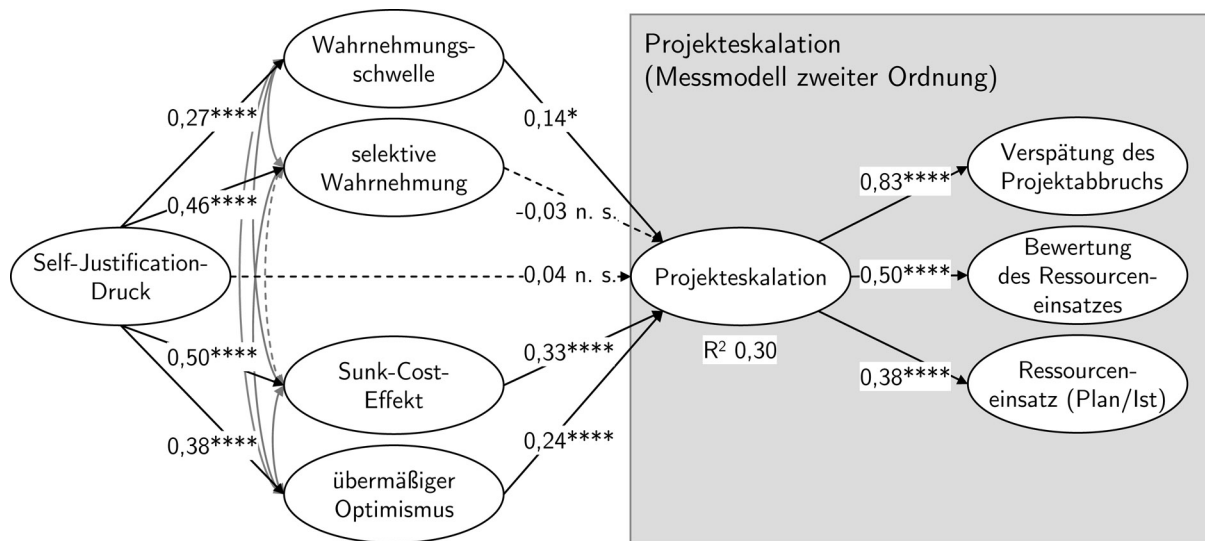
Zusätzlich wurde das Modell 3 auch ohne Kovarianzen zwischen den kognitiven Faktoren geschätzt. Im Vergleich zur Variante mit Kovarianzen fallen die Gütemaße etwas schlechter

²⁵³ Vgl. Festinger/Riecken/Schachter 1956, S. 26, Goethals 1992, S. 328, und Goethals/Reckman 1973.

²⁵⁴ Vgl. Abelson/Levi 1985, S. 276, Janis/Mann 1977, S. 317, und Cooper/Woo/Dunkelberg 1988, S. 106.

²⁵⁵ Vgl. Abbildung 4.6 auf der nächsten Seite.

²⁵⁶ Vgl. Abschnitt 4.2.3 auf Seite 117.



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df 2,105	< 3,00	0,55* standardisierter Pfadkoeffizient	* signifikant auf 0,1-Niveau
TLI 0,967	> 0,95	R ² Erklärungsgehalt	** signifikant auf 0,05-Niveau
CFI 0,973	> 0,95	Indikatoren, Fehlerterme und	*** signifikant auf 0,01-Niveau
RMSEA 0,040	< 0,08	Residuen sind zur Vereinfachung	**** signifikant auf 0,001-Niveau
SRMR 0,046	< 0,08	nicht mit abgebildet	n. s. nicht signifikant

Abbildung 4.6: Modell mit kognitiver Dissonanz (Modell 3)

aus ($\chi^2_{df} = 2,683$; TLI = 0,950; CFI = 0,958; RMSEA = 0,050 und SRMR = 0,068), sind jedoch immer noch akzeptabel. Die Pfadkoeffizienten, Signifikanzniveaus und R² verändern sich nur geringfügig, so dass die inhaltliche Interpretation identisch ist.

Die Konsequenzen dieser Ergebnisse für die Bewertung der kognitiven Ursachen der Projekteskalation werden in Abschnitt 4.4.1.5 auf Seite 160 diskutiert.

4.4.1.4 Exploratorische Modellselektion (Strukturmodell 4)

Nachdem drei Modelle konfirmatorisch überprüft wurden, soll ergänzend noch ein Modell exploratorisch ermittelt werden, das einen besonders guten Fit zu den Daten aufweist. Strukturgleichungsmodelle auch im Rahmen exploratorischer Ansätze zu ermitteln (auch als Post-hoc-Analyse oder Specification Search bezeichnet), findet in der Literatur seit einiger Zeit Beachtung.²⁵⁷

Die Modellselektion kann mit verschiedenen Verfahren durchgeführt werden. Weit verbreitet ist beispielsweise die iterative Modellselektion. Dabei ein Basismodell Schritt für Schritt „per

²⁵⁷ Vgl. Homburg/Dobratz 1991, S. 215.

Hand“ abgewandelt, um einen besseren Fit zu erzielen.²⁵⁸ Modifikationen eines Modells sollten dabei nur erfolgen, wenn sie sich auch theoretisch begründen lassen.²⁵⁹ Andernfalls besteht die Gefahr des Overfitting. Mit diesem Begriff werden Anpassungen bezeichnet, die lediglich für die vorliegende Stichprobe und nicht für die Grundgesamtheit gelten. In einem solchen Fall wird der Fit eines Modells zwar verbessert, der Grund dafür liegt jedoch ausschließlich in spezifischen Merkmalen der Stichprobe. Die Interpretation eines solchen Modells geht somit an der Realität (d. h. der Grundgesamtheit) vorbei.²⁶⁰

Aktuelle Computerprogramme ermöglichen auch weitgehend automatisierte exploratorische Analysen. So sind beispielsweise bei Amos 5.0 verschiedene Suchalgorithmen implementiert.²⁶¹ Auch die Ergebnisse einer solchen Suche sollten selbstverständlich im Hinblick auf die vorhandenen Theorien und empirischen Belege aus anderen Studien geprüft werden.

Für die Untersuchung der psychologischen Zusammenhänge bei der Eskalation von Projekten stellt sich die Frage, wie die kognitiven Faktoren untereinander wirken. Einige aus der Literatur abgeleitete Modelle wurden bereits in den vorherigen Abschnitten geprüft. Im Zuge einer Specification Search wird an dieser Stelle zusätzlich das Auftreten von Effekten zwischen allen fünf kognitiven Faktoren exploratorisch geprüft. Dazu muss bei dem Programm Amos zunächst angegeben werden, zwischen welchen Faktoren potentielle Wirkungen berücksichtigt werden sollen. Die gestrichelten Linien in Abbildung 4.7 geben die als optional gekennzeichneten Pfade wider, die im Rahmen der automatisierten exploratorischen Suche analysiert wurden. Für die Analyse wurde das „exhaustive“ Berechnungsverfahren gewählt, dass *alle möglichen* Kombinationen von Effekten durchrechnet. Für fünf Konstrukte handelt es sich dabei um 262.144 mögliche Modelle, die mit Amos berechnet wurden.

Zur Beurteilung der exploratorisch ermittelten Modelle kann das Browne-Cudeck-Criterion (BCC) herangezogen werden.²⁶² Als das beste Modell gilt dasjenige mit dem geringsten BCC. Werden die BCCs für alle geschätzten Modelle so transformiert, dass der kleinste BCC dem Wert null entspricht, so gilt die Regel, dass alle Modelle mit einem BCC von null bis zwei in ihrer Anpassungsgüte nicht signifikant unterschiedlich sind.²⁶³ Bei der vorliegenden Untersu-

²⁵⁸ Vgl. Homburg/Dobratz 1991.

²⁵⁹ Vgl. beispielsweise Cliff 1983, S. 123 f., Kaplan 1990, S. 153 und Raykov/Widaman 1995, S. 310 ff.

²⁶⁰ Vgl. Backhaus/Blechs Schmidt/Eisenbeiß 2006, S. 712.

²⁶¹ Bei den Suchalgorithmen wird zwischen „exhaustive“, „forward“, „backward“ und „stepwise“ unterschieden; vgl. Arbuckle 2006, S. 362.

²⁶² Vgl. Arbuckle 2006, S. 330 und 540, sowie Browne/Cudeck 1989.

²⁶³ Vgl. Burnham/Anderson 2002, S. 128.

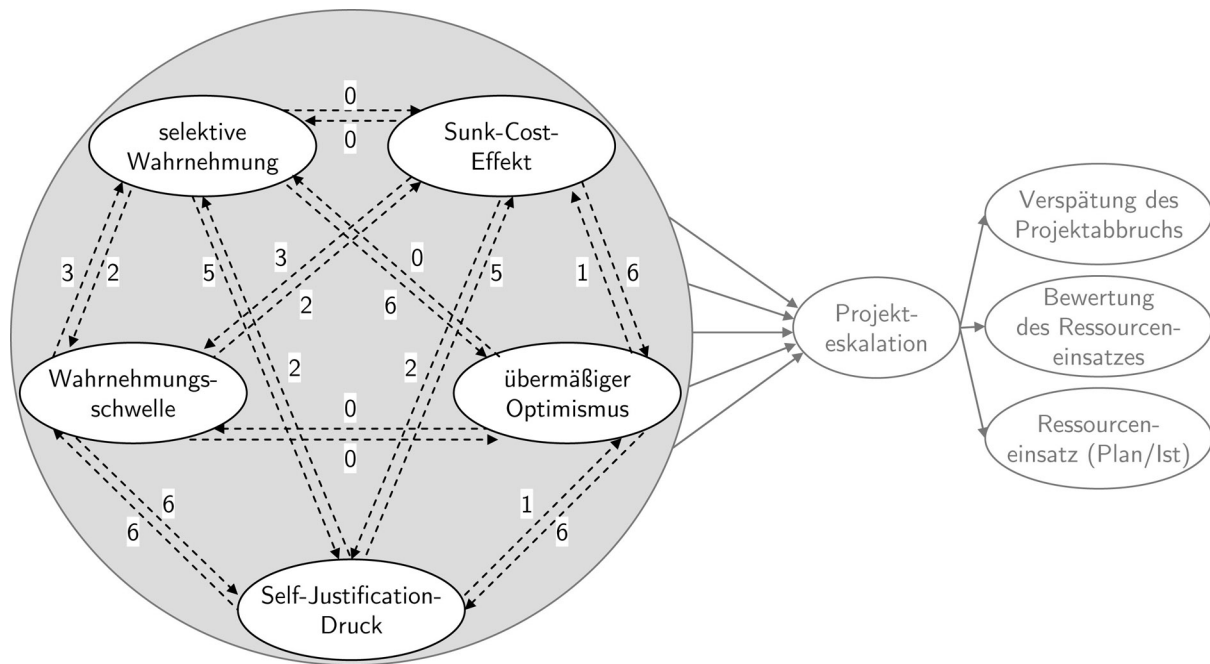


Abbildung 4.7: Specification Search der Wirkungen der kognitiven Faktoren

chung traf dies auf sieben stabile²⁶⁴ Modelle zu. Diese sieben Modelle mit einem BCC von unter 2,0 kommen demnach als mögliches bestes Modell in Frage.²⁶⁵

In Abbildung 4.7 ist neben jedem Pfad die Häufigkeit vermerkt, in der der Pfad bei den sieben ermittelten Modellen vorkam. Ein Pfad, der in keinem der Modelle vorgeschlagen wurde, trägt also die Ziffer null; ein Pfad, der in allen sieben Modellen vorgekommen wäre, hätte die Ziffer sieben erhalten. Mehrere Pfade tragen die Ziffer null. Die kann als Hinweis darauf gesehen werden, dass die Wirkung durch die empirischen Daten nicht bestätigt wird. Beispielsweise scheinen die selektive Wahrnehmung und der Sunk-Cost-Effekt voneinander unabhängig zu sein. Auch zwischen der Wahrnehmungsschwelle und dem übermäßigen Optimismus werden keine Wirkungen ermittelt. Fünf Pfade tragen die Ziffer sechs, wurden also in sechs von sieben Modellen vorgeschlagen. Diese werden im Folgenden kurz besprochen.

Eine häufig spezifizierte Wirkung übt die selektive Wahrnehmung auf den übermäßigen Optimismus aus. Dies ist intuitiv plausibel: Wenn ein Individuum positive, seine optimistische Sichtweise bestätigende Informationen stärker wahrnimmt, dann verstärken die selektiven Informationen den Optimismus und fördern eine unrealistische, übermäßig optimistische Sichtweise.

²⁶⁴ Im Rahmen der exploratorischen Suche können auch instabile Modelle auftreten (vgl. beispielsweise Arbuckle 2006, S. 351), die jedoch wenig aussagekräftig sind.

²⁶⁵ Vgl. Arbuckle 2006, S. 330.

Ein zweiter sehr häufig vorgeschlagener Effekt betrifft die Wirkung der Sunk Costs auf den übermäßigen Optimismus. Dies könnte durch die kognitive Dissonanztheorie erklärt werden. In ein scheiterndes Projekt viel investiert zu haben, ruft eine kognitive Dissonanz hervor. Diese kann dadurch verringert werden, dass die eigene Einstellung in Bezug auf die Erfolgswahrscheinlichkeit verändert wird, d. h. der übermäßige Optimismus gestärkt wird.

Für den übermäßigen Optimismus wird in sechs von sieben Modellen eine Wirkung auf den Self-Justification-Druck spezifiziert. Eine mögliche Erklärung dafür lautet wie folgt: Wenn der übermäßige Optimismus des Verantwortlichen zunimmt, steigt gleichzeitig die Erwartungshaltung an das Ergebnis und damit auch der Druck, dieser Erwartung gerecht zu werden. Dies gilt zunächst für das Selbstbild des Verantwortlichen. Kommuniziert er seine übermäßig optimistische Sicht an Mitarbeiter und Vorgesetzte, so steigt auch deren Erwartungshaltung. Um so stärker wäre dann das Image des Verantwortlichen gefährdet, wenn das Projekt doch abgebrochen würde.

Zwischen der Wahrnehmungsschwelle und dem Self-Justification-Druck werden in beide Richtungen Wirkungen durch die Specification Search nahegelegt. Allerdings ist dabei die Wirkung des Self-Justification-Drucks auf die Wahrnehmungsschwelle positiv, während der umgekehrte Effekt negativ ist, eine zunehmende Wahrnehmungsschwelle also zu einer Senkung des Self-Justification-Drucks führt.²⁶⁶ Beide Effekte sind sehr plausibel. Der Effekt des Self-Justification-Drucks auf die Wahrnehmungsschwelle lässt sich wiederum mit der kognitiven Dissonanztheorie erklären. Je weniger Informationen bei einem scheiternden Projekt wahrgenommen werden, desto weniger wird die Dissonanz verstärkt. Der abschwächende Effekt der Wahrnehmungsschwelle auf den Self-Justification-Druck ist intuitiv plausibel. Je weniger Informationen in einem scheiternden Projekt wahrgenommen werden, desto weniger Druck wird der Verantwortliche empfinden.

Abbildung 4.8 auf der nächsten Seite stellt das Modell mit dem kleinsten BCC (und somit dem besten Fit) dar, das im Rahmen der Specification Search ermittelt wurde. Wie zuvor ausgeführt, muss es mit Vorsicht interpretiert werden, da die Specification Search noch zu sechs weiteren Modellen führte, die ebenfalls als beste Modelle in Frage kommen, also nicht signifikant schlechter als Modell 4 zu den empirischen Daten passen. Das Modell 4 entspricht im Wesentlichen den Zusammenhängen in Abbildung 4.7 auf der vorherigen Seite. Zusätzlich erlaubt die genauere Analyse eines einzelnen Modells eine Betrachtung der standardisierten Pfadkoeffizienten.

²⁶⁶ Vgl. Abbildung 4.8 auf der nächsten Seite.

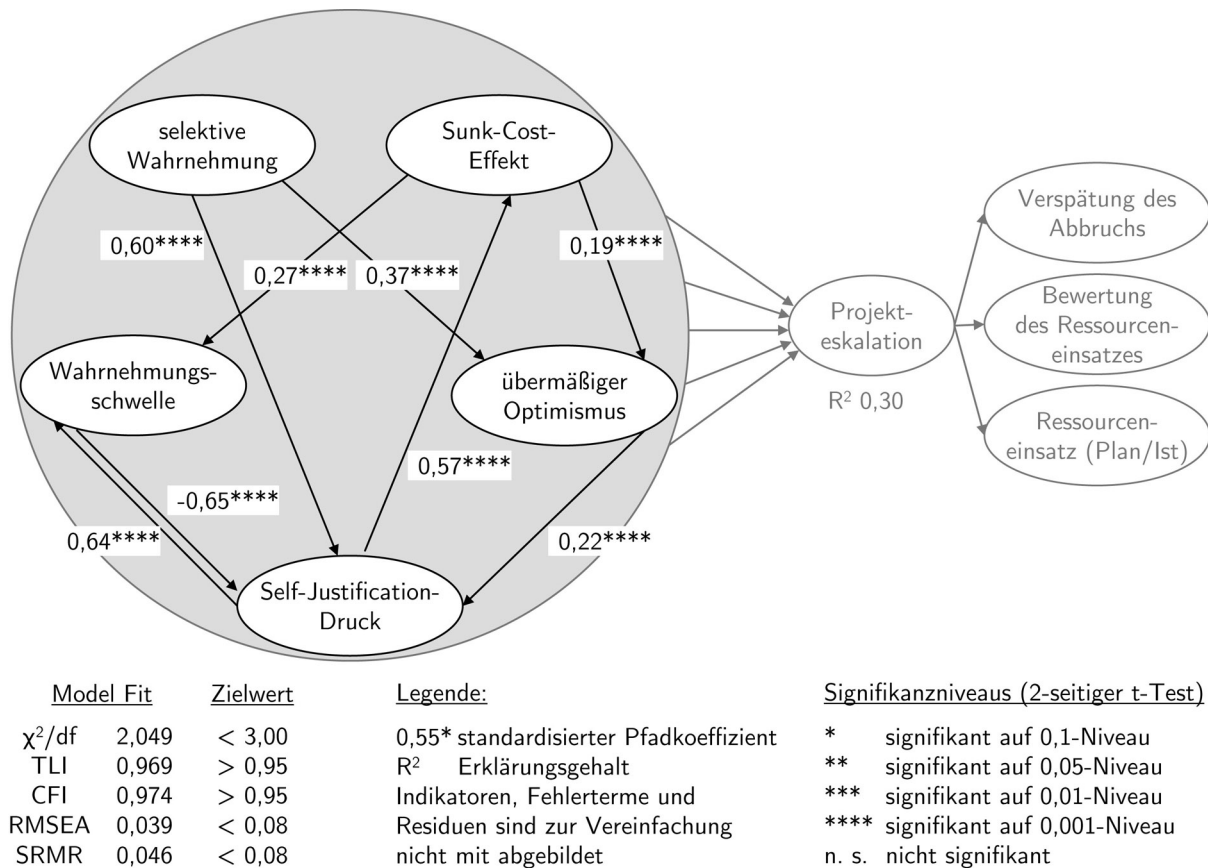


Abbildung 4.8: Eines von sieben exploratorisch ermittelten Modellen (Modell 4)

Dabei fällt auf, dass insbesondere die gegenseitigen Wirkungen von Wahrnehmungsschwelle und Self-Justification-Druck sehr hoch sind. Der Self-Justification-Druck scheint damit erheblichen Einfluss darauf zu haben, wie hoch die Schwelle ist, die neue Informationen überwinden müssen. Die negative Wirkung der Wahrnehmungsschwelle auf den Self-Justification-Druck könnte hierbei allerdings durch den Effekt der Suppression bei der statistischen Analyse hervorgerufen werden.²⁶⁷ Auf eine inhaltliche Interpretation soll deshalb verzichtet werden.

Ein sich verstärkender Kreislauf könnte gemäß Modell 4 zwischen Self-Justification-Druck, Sunk-Cost-Effekt und übermäßigem Optimismus bestehen. Ein höherer Self-Justification-Druck verstärkt den Sunk-Cost-Effekt, dieser bewirkt einen noch stärkeren übermäßigen Optimismus und dies wiederum führt zu einer Steigerung des Self-Justification-Drucks. Es liegt also gewissermaßen eine Eskalationsspirale vor. Falls die Wirkungen tatsächlich in dieser Art und Weise in der Realität auftreten, sollten zukünftige Forschungsprojekte untersuchen, wie derartige Kreisläufe durchbrochen werden können.

²⁶⁷ Vgl. Cohen et alii 2003, S. 77 f. und 457 f.

Der drittstärkste Effekt in Modell 4 geht von der selektiven Wahrnehmung auf den Self-Justification-Druck aus. Im Gegensatz zur Wahrnehmungsschwelle ist die Wirkung diesmal jedoch positiv gerichtet. Eine stärkere selektive Wahrnehmung führt zu höherem Self-Justification-Druck. Dies erscheint zunächst kontraintuitiv. möglicherweise liegt die Erklärung des Unterschiedes darin, dass bei einer höheren Wahrnehmungsschwelle auch weniger positive Informationen wahrgenommen werden. Dies würde generell zu weniger Aufmerksamkeit für die Ergebnisse des Projekts führen und damit den Druck senken. Durch die überwiegend positive Verzerrung der selektiven Wahrnehmung steigt hingegen möglicherweise die Erwartungshaltung. Diese Annahme deckt sich mit der signifikanten Wirkung selektiver Wahrnehmung auf übermäßigen Optimismus. Dabei ist sowohl die eigene Erwartungshaltung betroffen, als auch – wenn die selektiven Informationen in der selektiven Form weitergegeben werden – die Erwartungshaltung des Umfelds.

Die Argumentation für die Wirkung der selektiven Wahrnehmung lässt sich damit ähnlich führen wie die zum Effekt des übermäßigen Optimismus. In beiden Fällen wird das mentale Modell des Entscheidungsverantwortlichen und des Umfelds dahingehend verändert, dass einem erfolgreichen Ausgang des Projekts eine höhere Wahrscheinlichkeit zugesprochen wird und damit der Projektverantwortliche unter noch größerem Druck steht, dieses Ergebnis tatsächlich auch zu liefern.

Bei allen Argumentationen in diesem Abschnitt handelt es sich jedoch vorerst nur um Hypothesen, die zukünftiger empirischer Prüfung bedürfen.

4.4.1.5 Zusammenführung der Ergebnisse der vier betrachteten Modelle

Im Folgenden sollen die bisherigen Ergebnisse kurz rekapituliert und gleichzeitig die Resultate der Parameterschätzungen für Modell 1 bis 4 in Bezug zueinander gesetzt werden.

(1) Mit Modell 1 konnte gezeigt werden, dass die kognitiven Faktoren in realen Unternehmensprojekten nicht vollständig unabhängig voneinander auftreten. Der Versuch, ein solches Modell zu spezifizieren, führt zu unzureichenden Gütemaßen.

(2) Werden hingegen, wie in Modell 2, Kovarianzen zwischen den kognitiven Faktoren zugelassen, ergeben sich sehr zufriedenstellende Gütemaße. Von den fünf betrachteten kognitiven Faktoren zeigen drei eine kontextübergreifende Wirkung auf die Projekteskalation: die Wahrnehmungsschwelle, der Sunk-Cost-Effekt und der übermäßige Optimismus. Diese drei Faktoren können 30 Prozent der Varianz der Projekteskalation erklären. Die zwei Faktoren selektive

Wahrnehmung und Self-Justification-Druck weisen keine direkte Wirkung auf die Projekteskalation auf.

(3) Aus der Schätzung von Modell 3 sowie aus der exploratorischen Analyse und dem daraus entwickelten Modell 4 lassen sich mehrere Schlussfolgerungen ableiten, mit denen die Resultate der konfirmatorischen Modellprüfung von Modell 1 und 2 um zusätzliche Erkenntnisse ergänzt werden. So unterstützen die Parameterschätzungen die These, dass in realen Projekten komplexe Wirkungsbeziehungen zwischen den kognitiven Faktoren bestehen. Diese Wirkungen wurden in der Literatur bisher kaum untersucht. Das auf der Basis der kognitiven Dissonanztheorie spezifizierte Modell 3 zeigt, dass die Daten beispielsweise einen guten Fit aufweisen, wenn der Self-Justification-Druck als Prädiktor für die vier anderen kognitiven Faktoren spezifiziert wird.

(4) Die exploratorischen Befunde und Modell 4 geben zusätzliche Anregungen für zukünftige Forschungen. Besondere Beachtung verdient dabei, dass zwischen selektiver Wahrnehmung und Sunk-Cost-Effekt sowie zwischen Wahrnehmungsschwelle und übermäßigem Optimismus kein direkter Zusammenhang zu bestehen scheint. Ebenfalls auffällig und für zukünftige Untersuchungen interessant ist die wechselseitige Wirkung von Wahrnehmungsschwelle und Self-Justification-Druck. Ein sich verstärkender Kreislauf könnte gemäß Modell 4 zwischen Self-Justification-Druck, Sunk-Cost-Effekt und übermäßigem Optimismus bestehen. Sollte ein solcher Zusammenhang in zukünftigen Studien bestätigt werden können, wäre es von großer Bedeutung, Wege zu finden, diese Eskalationsspirale zu durchbrechen.

(5) Das ermittelte Wirkungsgeflecht zwischen den Faktoren wirft im Hinblick auf die selektive Wahrnehmung und den Self-Justification-Druck ein neues Licht auf die konfirmatorischen Ergebnisse von Modell 1 und 2. In den Schätzungen von Modell 1 und 2 sind die Wirkungen dieser beiden Faktoren auf die Projekteskalation nicht signifikant. Für sich genommen legen diese Resultate den Schluss nahe, dass selektive Wahrnehmung und Self-Justification-Druck verhältnismäßig unbedeutend für die Erklärung der Projekteskalation sind und deshalb in diesem Zusammenhang weder weiter untersucht noch explizit in der Praxis berücksichtigt werden müssen. Die exploratorischen Resultate hingegen zeigen, dass beide Faktoren signifikante Wirkungen auf andere kognitive Faktoren ausüben und somit indirekt einen starken Einfluss auf die Projekteskalation haben können. Entgegen den konfirmatorisch basierten Schlussfolgerungen sollten sie dementsprechend durchaus weiter erforscht und in der Praxis berücksichtigt werden. Somit kommt auch den Controlling-Maßnahmen, die helfen, selektive Wahrnehmung und Self-Justification-Druck abzusenken, eine wichtige Bedeutung zu.

Zusammenfassend ist Folgendes festzuhalten: Die Relevanz der in Zayers Modell integrierten kognitiven Faktoren für die Projekteskalation wird durch die empirischen Daten aus realen Un-

ternehmenssituationen gestützt. Um einen angemessenen Fit zwischen Modell und Daten zu erreichen, musste die vereinfachende Annahme, dass die kognitiven Faktoren unabhängig voneinander seien, aufgehoben werden. Einige mögliche Wirkungsmechanismen zwischen den kognitiven Faktoren wurden exploratorisch ermittelt. Um diese genauer zu erforschen oder ggf. zu falsifizieren, sollten zukünftige Studien zunächst die psychologische Literatur gezielt nach Erklärungen und Theorien für die gefundenen Wirkungen durchsuchen. Basierend darauf, sollten mit erneuten empirischen Erhebungen konfirmatorische Prüfungen erfolgen.

4.4.2 Analyse der moderierenden Effekte

Das konzeptionelle Modell von Zayer beinhaltet mehrere moderierende Effekte.²⁶⁸ Zu ihrer Prüfung werden in den folgenden Abschnitten Gruppenvergleichsrechnungen durchgeführt.²⁶⁹ Dabei wird zuerst für jedes moderierende Konstrukt die Summe aller Indikatoren gebildet. Für den so entstandenen Index wird anschließend ein Trennwert so festgelegt, dass das Sample in zwei etwa gleich große Hälften – eine mit niedriger und eine mit hoher Ausprägung – aufgeteilt wird (Split-Half-Verfahren). Die Modelle für beide Gruppen werden simultan geschätzt, wobei der zu untersuchende Pfad einmal restringiert (d. h. für beide Gruppen als gleich festgelegt) und einmal frei geschätzt wird. Schließlich wird mit einem χ^2 -Differenztest²⁷⁰ geprüft, ob sich das restringierte und das frei geschätzte Modell in ihrem Fit signifikant unterscheiden. Ist das restringierte Modell signifikant schlechter, unterscheiden sich die beiden Gruppen, womit der moderierende Effekt als bestätigt gilt. Zusätzlich zu der Aufteilung in zwei gleich große Hälften wird das Sample auch in drei Gruppen unterteilt (niedrig, mittel, hoch) und das Drittel mit der niedrigen Ausprägung mit dem Drittel mit der hohen Ausprägung verglichen.

Für die folgenden Analysen der moderierenden Effekte wurden zur Komplexitätsreduzierung *nicht* die Gesamtmodelle mit allen kognitiven Faktoren herangezogen, sondern jeweils nur der betreffende einzelne kognitive Faktor und die Projekteskalation modelliert.

²⁶⁸ Vgl. Abschnitt 2.3.5 auf Seite 41.

²⁶⁹ Vgl. Arbuckle 2006, S. 367 ff., und Byrne 2001, S. 173 ff.

²⁷⁰ Vgl. beispielsweise Schermelleh-Engel/Moosbrugger/Müller 2003, S. 33 ff.

4.4.2.1 Zeitpunkt des ersten Problemauftretens

Es wurde hypothetisiert, dass der Zeitpunkt des ersten Problemauftretens als moderierender Faktor die Wirkung der Wahrnehmungsschwelle auf die Projekteskalation beeinflusst.²⁷¹ Dabei wurde angenommen, dass die Projekteskalation umso geringer ausfällt, je mehr negative Informationen zu einem *frühen* Zeitpunkt auftreten. Wenn also die Informationen früh auftreten, wird ein positiver Zusammenhang zwischen Wahrnehmungsschwelle und Projekteskalation vermutet: Je höher die Wahrnehmungsschwelle, desto größer die Projekteskalation. Treten die Probleme erst *spät* auf, so wurde angenommen, dass ein negativer Zusammenhang zwischen Wahrnehmungsschwelle und Projekteskalation auftritt: Je höher die Wahrnehmungsschwelle, desto geringer die Projekteskalation. Es handelt sich somit um eine negative Moderation.

Die Prüfung dieser Hypothesen durch einen Gruppenvergleich wird in Abbildung 4.9 auf der nächsten Seite dargestellt. Die Pfadkoeffizienten in Klammern entsprechen der unstandardisierten Lösung, die Pfadkoeffizienten ohne Klammern geben den standardisierten Schätzwert wieder.

Die beiden Modelle unterscheiden sich um einen Freiheitsgrad, die Differenz des χ^2 -Werts beträgt 0,651. In der üblichen Kurzschreibweise wird dies als $\chi^2_{(1)} = 0,651$ notiert. Der Unterschied ist nicht signifikant ($p = 0,42$). Der Effekt wird auch dann nicht signifikant, wenn das Gesamtsample in drei Gruppen (früher, mittlerer und später Zeitpunkt) aufgeteilt wird und die Gruppe „früher Zeitpunkt“ mit der Gruppe „später Zeitpunkt“ verglichen wird ($\chi^2_{(1)} = 0,199$; $p = 0,66$; vgl. Abbildung 4.9 auf der nächsten Seite). Die Hypothese eines moderierenden Effekts des Zeitpunkts der ersten Problemwahrnehmung (H_6)²⁷² bezüglich der Wirkung der Wahrnehmungsschwelle auf die Projekteskalation wird damit abgelehnt.

Eine mögliche Erklärung könnte darin liegen, dass in dem Experiment, das der Hypothese H_6 zu Grunde liegt,²⁷³ die Gruppe, die sich im signifikant von den anderen unterschied, die Kosteninformationen bereits vor Beginn ihrer Investitionen erhielt. Ein großer Teil des Effekts könnte somit auf die Personen zurückzuführen sein, die auf der Basis der Kosteninformationen gar nicht erst mit dem Investitionsprozess begonnen haben.²⁷⁴ Demzufolge ist die Übertragbarkeit auf bereits laufende, scheiternde Projekte eingeschränkt. Um zu prüfen, ob der von Brockner et alii berichtete Effekt in der Unternehmenspraxis nachweisbar ist, wäre es deshalb sinnvoll, wenn

²⁷¹ Vgl. Abschnitt 2.3.5.1 auf Seite 41.

²⁷² Vgl. Abbildung 2.3 auf Seite 28.

²⁷³ Vgl. Brockner et alii 1982, S. 251.

²⁷⁴ Vgl. Brockner et alii 1982, S. 253.

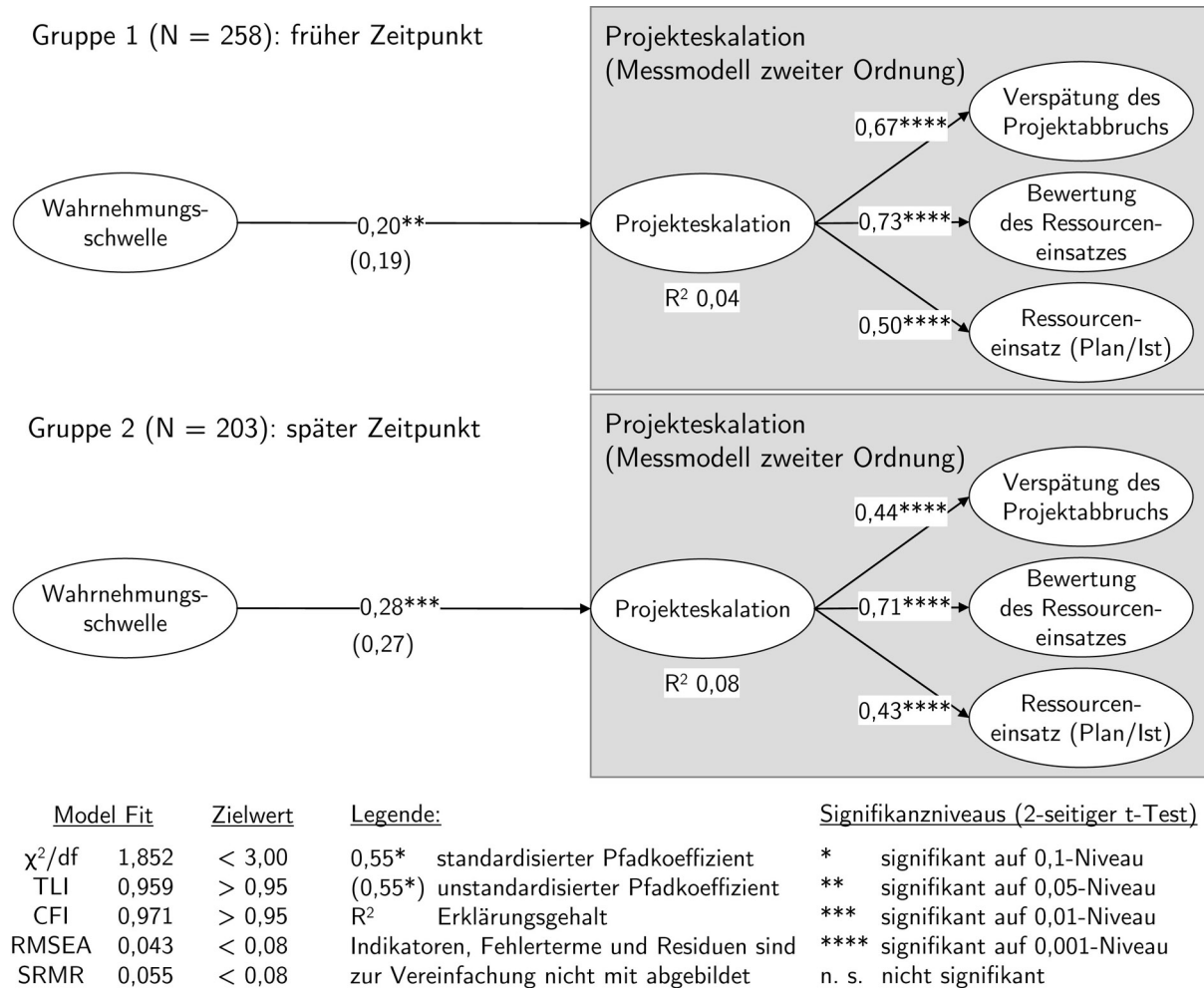


Abbildung 4.9: Zeitpunkt des ersten Problemauftretens als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)

nachfolgende Studien zusätzlich zu gescheiterten Projekten auch ein Teilsample untersuchen, in dem Projekte schon während der Planungsphase verworfen, also gar nicht erst begonnen, wurden.

4.4.2.2 Nicht hinterfragter Freiraum

Der nicht hinterfragte Freiraum wirkt sich entsprechend H_7 moderierend auf die Beziehung zwischen Self-Justification-Druck und Projekteskalation aus.²⁷⁵ Die Prüfung dieser Hypothese durch einen Gruppenvergleich wird in Abbildung 4.10 auf der nächsten Seite dargestellt.

²⁷⁵ Vgl. Abschnitt 2.3.5.2 auf Seite 42 und Abbildung 2.3 auf Seite 28.

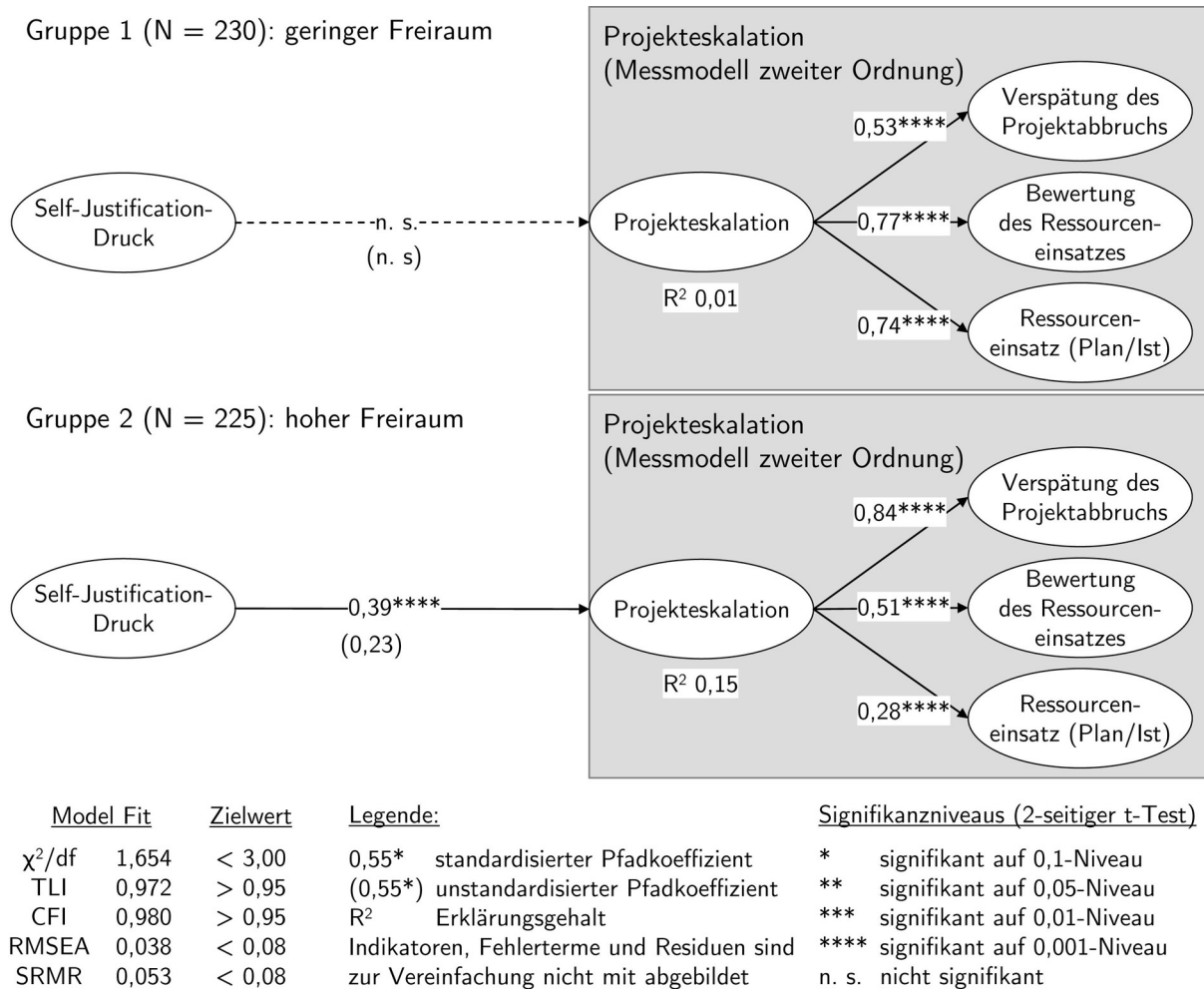


Abbildung 4.10: Nicht hinterfragter Freiraum als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)

Wenn das Sample in zwei Gruppen aufgeteilt wird, lässt sich der moderierende Effekt des nicht hinterfragten Freiraums nicht statistisch signifikant bestätigen ($\chi^2_{(1)} = 0,568$; $p = 0,45$). Teilt man das Sample jedoch in drei Gruppen (niedriger, mittlerer und hoher Freiraum) ein und vergleicht dann niedrigen und hohen Freiraum, so wird der moderierende Effekt deutlich signifikant ($\chi^2_{(1)} = 6,296$; $p = 0,012$; vgl. Abbildung 4.10). Für die Gruppe mit geringem Freiraum (N = 230) tritt keine signifikante Wirkung des Self-Justification-Drucks auf die Projekteskalation auf. Die Varianz der abhängigen Variablen Projekteskalation wird nur zu einem Prozent erklärt ($R^2 = 0,01$). Für die Gruppe mit großem nicht hinterfragtem Freiraum (N = 225) beträgt der standardisierte Pfadkoeffizient 0,39, der Effekt ist auf einem Niveau von 0,001 signifikant.

Als Resultat der statistischen Tests kann festgehalten werden, dass Gruppe 2 einen signifikant höheren nicht hinterfragten Freiraum besitzt und in dieser Situation der Self-Justification-Effekt signifikant auf die Projekteskalation wirkt. Dies stützt die in Abschnitt 2.3.5.2 auf Seite 42

dargestellte These, dass Menschen dem Self-Justification-Druck nachgeben und ein Projekt zum Selbstschutz fortführen, wenn sie dank eines großen nicht hinterfragten Freiraums die Möglichkeit dazu haben. Die Hypothese eines moderierenden Effekts des nicht hinterfragten Freiraums (H_7) bezüglich der Wirkung des Self-Justification-Drucks auf die Projekteskalation wird damit gestützt.

4.4.2.3 Overconfidence

Gemäß H_8 wird vermutet, dass sich Overconfidence moderierend auf die Beziehung zwischen übermäßigem Optimismus und Projekteskalation auswirkt.²⁷⁶ Die Prüfung dieser Hypothese durch einen Gruppenvergleich wird in Abbildung 4.11 auf der nächsten Seite dargestellt.

Weder beim Split-Half-Verfahren ($\chi^2_{(1)} = 0,236$; $p = 0,63$) noch beim Vergleich des niedrigsten mit dem höchsten Drittel treten signifikante Modellunterschiede auf ($\chi^2_{(1)} = 0,079$; $p = 0,79$; vgl. Abbildung 4.11). Eine moderierende Wirkung von Overconfidence bezüglich der Wirkung von übermäßigem Optimismus auf die Projekteskalation wird demzufolge nicht bestätigt. H_8 muss somit verworfen werden.

Dies bedeutet jedoch nicht, dass Overconfidence zum Verständnis des übermäßigem Optimismus keinen Beitrag leistet. Wie aus der konfirmatorischen Faktorenanalyse hervorgeht,²⁷⁷ zeigen die Daten eine signifikante positive Korrelation in Höhe von 0,35 zwischen übermäßigem Optimismus und Overconfidence. Welche Kausaleffekte diesen Zusammenhang hervorrufen, sollte in zukünftigen Studien weiter erforscht werden.

4.4.3 Beurteilung des Akteursmodells

In Abschnitt 2.1.3 wurde dargelegt, wie die kognitiven Eskalationsfaktoren mit Hilfe der Methode der abnehmenden Abstraktion mit der klassischen ökonomischen Theorie verknüpft werden können. In diesem Abschnitt soll kurz erläutert werden, ob die vorgenommenen Erweiterungen sich in den empirischen Daten wiederfinden lassen und wie groß der Erklärungsgehalt für die Projekteskalation ist. Anschließend wird der Nutzen der Methode der abnehmenden Abstraktion kurz diskutiert.

²⁷⁶ Vgl. Abschnitt 2.3.5.3 auf Seite 43 und Abbildung 2.3 auf Seite 28.

²⁷⁷ Vgl. Tabelle 4.23 auf Seite 145.

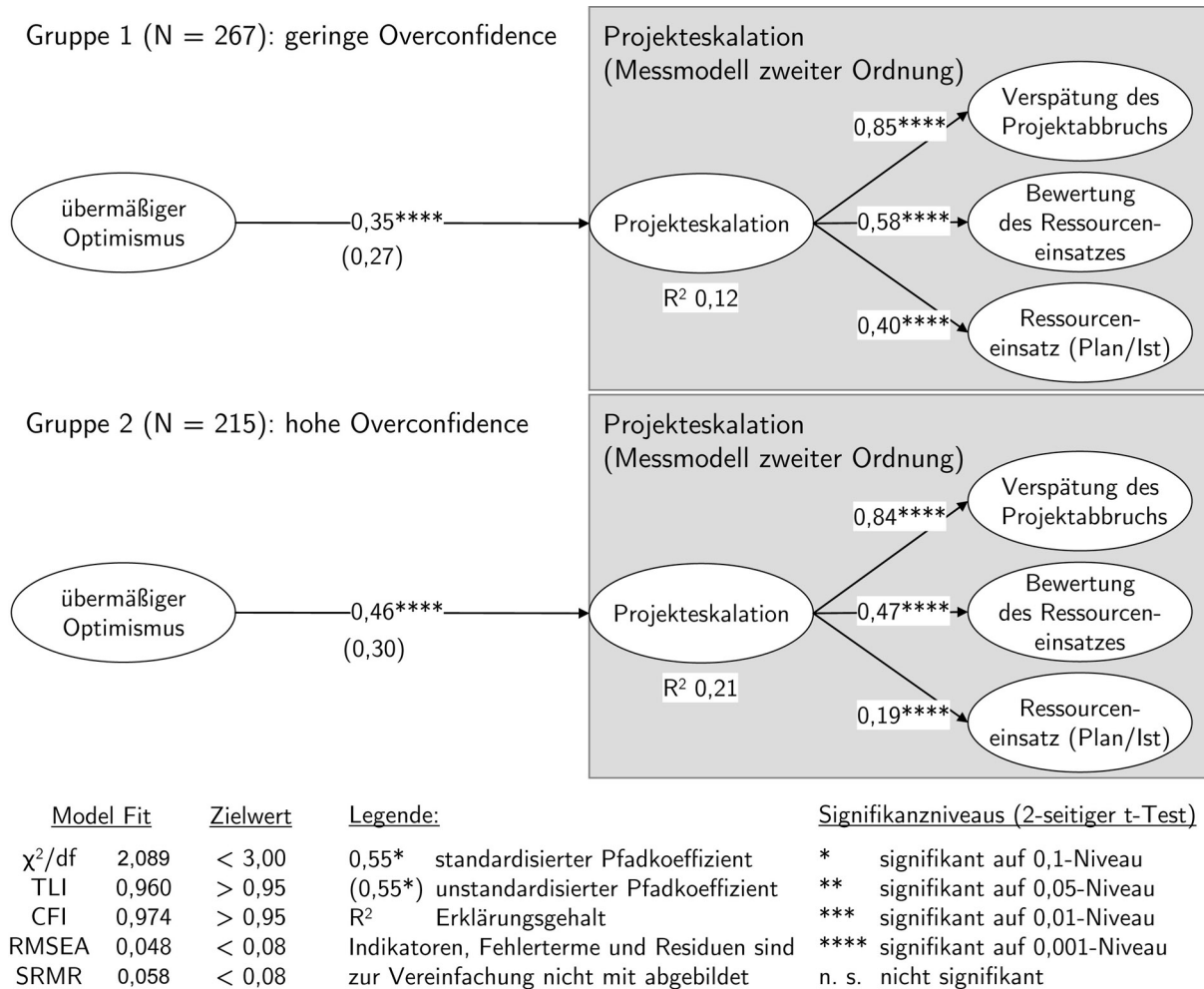


Abbildung 4.11: Overconfidence als moderierender Faktor (Vergleich zwischen niedrigstem Drittel und höchstem Drittel des Datensatzes)

Die in den letzten Abschnitten betrachteten Auswirkungen der kognitiven Faktoren weisen eine Gemeinsamkeit auf: Entsprechend der klassischen ökonomischen Theorie dürften sie nicht auftreten. Für die erfolgte Messung bedeutet dies, dass die Antworten bei Gültigkeit des homo oeconomicus²⁷⁸ auf der siebenstufigen Skala von „0 = nie“ bis „6 = immer“ stets „0 = nie“ hätten lauten müssen. Dies ist in der empirischen Messung nicht der Fall. Die Durchschnittswerte der Antworten zu allen Fragen sind deutlich von null verschieden (siehe Tabelle A.2 auf Seite 236 im Anhang). Demzufolge stehen die empirischen Daten im Einklang mit den vorgenommenen *Erweiterungen* des ökonomischen Modells. Bei keiner Frage wird die strenge Annahme der klassischen Ökonomie bzw. des homo oeconomicus bestätigt. Stattdessen geben die Teilnehmer mehr oder weniger häufig Abweichungen von den Annahmen einer vollständigen und ra-

²⁷⁸ Vgl. dafür die Ausführungen zum RREEMM in Abschnitt 2.1.3.

tionalen Wahrnehmung und Beurteilung zu Protokoll. Die empirischen Daten weisen also einen Fit mit allen vorgenommenen Erweiterungen auf.

Nachdem festgestellt wurde, dass die Erweiterungen des klassischen ökonomischen Modells sich in den empirischen Daten wiederfinden, stellt sich die Frage, ob die Erweiterungen auch einen Beitrag zur Erklärung der Projekteskalation leisten. Dies wurde in den Analysen der Strukturmodelle bestätigt.²⁷⁹ Die Erweiterungen der Resourceful-Annahme durch die Wahrnehmungsschwelle, die Erweiterung der Evaluating-Annahme durch die Sunk Costs und die Erweiterung der Expecting-Annahme durch den übermäßigen Optimismus können signifikant zur Erklärung der Projekteskalation beitragen. Die Erweiterung der Evaluating-Annahme durch Self-Justification trägt signifikant zur Erklärung der Projekteskalation bei, wenn der nicht hinterfragte Freiraum groß ist. Lediglich die Erweiterungen der selektiven Wahrnehmung und der Overconfidence konnten für die untersuchten Daten keinen *direkten* signifikanten Erklärungsbeitrag leisten. Wie jedoch die exploratorischen Analysen gezeigt haben, ist davon auszugehen, dass die selektive Wahrnehmung auf andere kognitive Faktoren wirkt und somit *indirekt* einen Erklärungsbeitrag für die Projekteskalation aufweist.²⁸⁰ Tabelle 4.25 fasst die besprochenen Ergebnisse noch einmal zusammen.

Tabelle 4.25: Empirische Beurteilung der Erweiterungen des ökonomischen Modells

Annahme	Erweiterung	Fit	Empirische Daten zeigen Wirkung des kognitiven Faktors auf die Projekteskalation
Resourceful	(1) Wahrnehmungsschwellen	ja	ja, kontextübergreifend
	(2) selektive Wahrnehmung	ja	keine direkte Wirkung, aber kontextübergreifende indirekte Wirkung über andere Faktoren
Restricted	keine Erweiterung		
Evaluating	(3) Self-Justification	ja	ja, direkter kontextspezifischer Effekt bei großem nicht hinterfragtem Entscheidungsfreiraum sowie indirekter kontextübergreifender Effekt über andere Faktoren
	(4) Sunk-Cost-Effekt	ja	ja, kontextübergreifend
Expecting	(5) übermäßiger Optimismus	ja	ja, kontextübergreifend
	(6) Overconfidence	ja	nein
Maximizing	keine Erweiterung		

Fit = Empirische Daten stehen im Einklang mit der Erweiterung, d. h. das Auftreten der kognitiven Begrenzung wurde in den Messungen sichtbar, da die Mittelwerte deutlich von null verschieden sind.

²⁷⁹ Vgl. Kapitel 4.4.1 auf Seite 148.

²⁸⁰ Das Phänomen der Overconfidence wurde in der exploratorischen Analyse nicht mit untersucht, da es im konzeptionellen Modell lediglich als moderierender Faktor auftritt. Es ist aber nicht auszuschließen, dass auch die Overconfidence indirekt über andere kognitive Faktoren auf die Projekteskalation wirkt.

Der Nutzen der Methode der abnehmenden Abstraktion besteht darin, dass die empirischen Befunde und das klassische ökonomische Modell zueinander in Bezug gesetzt werden können. Dies bringt sowohl für die empirischen Ergebnisse als auch für das Verständnis des klassischen ökonomischen Modells Vorteile. Die empirischen Ergebnisse stehen dabei nicht isoliert nebeneinander, sondern sind durch wenige, klar genannte Annahmen mit dem Modell des homo oeconomicus verknüpft. Für das Modell des homo oeconomicus wird deutlich, dass mit wenigen Erweiterungen ein deutlich höherer Erklärungsgehalt für das zu untersuchende Phänomen erzielt werden kann.

Die statistischen Analysen zeigen, dass insbesondere die Erweiterungen der Evaluating-Annahme eine sehr starke Verbesserung des Erklärungsgehalts bieten. Aber auch die Erweiterungen der Resourceful- und der Expecting-Annahme weisen einen verbesserten Erklärungsgehalt im Vergleich zum klassischen ökonomischen Modell auf. Die Erweiterung der Expecting-Annahme durch die Overconfidence erwies sich (zumindest mit den durchgeführten Tests) als verzichtbar für die Erklärung der Projekteskalation.

Die Methode der abnehmenden Abstraktion und die anschließende empirische Prüfung können somit dazu beitragen, im Prozess der Theoriebildung der Forderung nachzukommen, Hypothesen mit größerem Erklärungsgehalt von Hypothesen mit geringerer Relevanz zu trennen.²⁸¹ Die Methode der abnehmenden Abstraktion als Bindeglied zwischen theoriegesteuerter und daten-gesteuerter Analyse²⁸² hat sich somit im Rahmen dieser Untersuchung bewährt.

4.5 Analyse der Maßnahmen des verhaltensorientierten Controllings

4.5.1 Wirkung der Maßnahmen auf die kognitiven Faktoren

Nachdem in den vorherigen Abschnitten die kognitiven Faktoren genauer untersucht wurden, liegt der Schwerpunkt dieses Abschnitts auf den Maßnahmen zur Reduzierung dieser Faktoren.

²⁸¹ Vgl. für diese Forderung Knight 1971, S. 4, und Friedman 1971, S. 23 ff.

²⁸² Vgl. Lindenberg 1991, S. 37.

Diese Perspektive ist in der Forschung lange vernachlässigt worden: “Much of the research in this area has focused on the determinants of escalation, and little attention has yet been given to procedures that might help people avoid the escalation trap. This is unfortunate because [...] deescalation research could potentially aid managers in avoiding the overcommitment of resources.”²⁸³

Für die folgenden Analysen wurde jeweils die Wirkung einer einzelnen Maßnahme²⁸⁴ auf den ihr zugeordneten kognitiven Faktor ermittelt, wobei der kognitive Faktor als Element des gesamten Eskalationsmodells (gemäß dem Strukturmodell in Abbildung 4.5 auf Seite 150) modelliert wurde.

4.5.1.1 Maßnahmen zur Reduzierung der Wahrnehmungsschwelle

Zur Reduzierung der Wahrnehmungsschwelle wurden sechs Maßnahmen aus der Literatur abgeleitet.²⁸⁵ Diese Maßnahmen wurden in der empirischen Erhebung auf ihre Einsatzhäufigkeit abgefragt. Die Antwortmöglichkeiten lagen (für diese und alle anderen Maßnahmen) auf einer siebenstufigen Skala von „0 = nie“ bis „6 = immer“. Die Durchschnittswerte der absoluten Einsatzhäufigkeit werden in Abschnitt 4.5.3 auf Seite 178 vorgestellt. Für die Auswertungen in diesem und den folgenden Abschnitten wurden die Kovarianzen der Einsatzhäufigkeit der Maßnahmen mit den Ausprägungen der kognitiven Faktoren untersucht, um Rückschlüsse auf die Wirksamkeit der Maßnahmen zu ziehen.²⁸⁶ Der Wortlaut der jeweiligen Maßnahme, die Kurz-

²⁸³ Simonson/Staw 1992, S. 419.

²⁸⁴ Der Grund dieser Vorgehensweise ist das Problem der Suppression, das zu einer Verzerrung der Regressionsgewichte führen kann, wenn mehrere Maßnahmen gleichzeitig geprüft werden; vgl. Cohen et alii 2003, S. 77 f. und 457 f.

²⁸⁵ Vgl. Abschnitt 2.3.2.1 auf Seite 30.

²⁸⁶ Diese Analysen basieren auf der vereinfachten Annahme eines linearen Zusammenhangs zwischen Maßnahme und kognitivem Faktor. Dies ist für den untersuchten Sachverhalt aus mehreren Gründen plausibel. Zum einen kann angenommen werden, dass bei der Mehrzahl der Projekte eher zu wenig Controllingmaßnahmen eingesetzt wurden als zu viel, da es sich um gescheiterte, eskalierte Projekte handelt. Trifft diese Annahme zu, liegt die Einsatzhäufigkeit unter dem Optimum, so dass von einem quasi-linearen Zusammenhang in dem beobachteten Abschnitt auszugehen ist. Zum anderen beträgt das Maximum der Antwortskala „immer“, was im Zusammenhang mit den Formulierungen der Maßnahmen als „immer in entsprechenden Situationen“ und nicht als „ständig“ im Sinne von zu viel zu verstehen ist. Beispielsweise ist für die Aussage „Es wurde regelmäßig (z. B. an jedem Meilenstein) aktiv über die Fortführung des Projekts entschieden“ die Antwort „immer“ als „zu jedem Meilenstein“ zu verstehen. Auch wenn diese Antwortskala bewusst gewählt wurde, weil sie als am besten geeignete Variante erscheint, wäre es sinnvoll, im Rahmen einer zukünftigen Studie zu prüfen, inwiefern sich die Ergebnisse verändern, wenn eine andere Antwortskala verwendet wird (beispielsweise eine Skala mit konkreten Zeitangaben wie „täglich“, „wöchentlich“ oder „monatlich“).

bezeichnung sowie die statistisch ermittelte Wirkung auf die Wahrnehmungsschwelle werden in Tabelle 4.26 dargestellt.

Tabelle 4.26: Maßnahmen zur Überwindung der Wahrnehmungsschwelle

Kurzbezeichnung	Formulierung	Wirkung
Meilensteine	Es wurden Meilensteine festgelegt, d. h. messbare Zwischenziele mit Deadline (Termin).	-0,08**
Genehmigung	Das Projekt musste jeweils vor Beginn der nächsten Teilphase eine Genehmigung erhalten, bevor es fortgesetzt wurde.	n. s.
aktive Entscheidung	Es wurde regelmäßig (z. B. an jedem Meilenstein) aktiv über die Fortführung des Projekts entschieden.	n. s.
Berichte	Es wurden Projektfortschrittsberichte (z. B. mit Kennzahlen über Abweichungen bei Kosten, Zeit, Qualität) für mich erstellt.	n. s.
eindeutige Kriterien	Es gab eindeutige Beurteilungskriterien, nach denen die Projektinformationen bewertet werden konnten (z. B. Zielwerte oder Benchmarkingstudien).	n. s.
Quantifizierung	Die Informationen über den Projektverlauf wurden für mich quantifiziert (d. h., mir lagen konkrete Zahlenwerte vor).	-0,11***

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

Die Maßnahmen lassen sich in zwei Gruppen aufteilen.²⁸⁷ Die ersten drei beziehen sich vor allem auf das Erreichen bewusster Entscheidungen über den Projektfortgang. Die Maßnahmen vier bis sechs zielen auf eine Erhöhung der Salience der relevanten Informationen ab.

Von den sechs Maßnahmen wirken zwei signifikant in der hypothetisierten Weise: Meilensteine und Quantifizierung. Zur Reduzierung der Wahrnehmungsschwelle scheinen im branchenübergreifenden Durchschnitt also die Verwendung von Meilensteinen und die Quantifizierung von Informationen besonders geeignet zu sein. Die Hypothesen H_{M1} und H_{M6} werden damit bestätigt.²⁸⁸ Alle anderen Maßnahmen weisen keine signifikante Wirkung auf. H_{M2} bis H_{M5} müssen demnach abgelehnt werden.

4.5.1.2 Maßnahmen zur Reduzierung der selektiven Wahrnehmung

In Abschnitt 2.3.2.2 auf Seite 34 wurden zwei Maßnahmen identifiziert, die zur Reduzierung der selektiven Wahrnehmung beitragen können. Diese Maßnahmen werden in Tabelle 4.27 auf der nächsten Seite aufgelistet.

²⁸⁷ Vgl. Abschnitt 2.3.2.1 auf Seite 30.

²⁸⁸ Vgl. Abschnitt 2.3.2.1 auf Seite 30 und Abbildung 2.3 auf Seite 28.

Tabelle 4.27: Maßnahmen zur Überwindung der selektiven Wahrnehmung

Kurzbezeichnung	Formulierung	Wirkung
Gefahrenhinweis	Ich wurde explizit auf die Gefahr hingewiesen, dass Projekte häufig zu spät abgebrochen werden (z. B. anhand von Daten vergangener Projekte).	n. s.
Abbruchhinweis	Mir wurde verdeutlicht, dass ich das Projekt abbrechen muss, wenn die Ziele nicht mehr erreicht werden können.	n. s.

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

Beide Maßnahmen erweisen sich in der empirischen Prüfung für das untersuchte Sample als nicht signifikant in ihrer Wirkung auf die selektive Wahrnehmung. H_{M7} und H_{M8} müssen deshalb verworfen werden.²⁸⁹

4.5.1.3 Maßnahmen zur Reduzierung des Self-Justification-Drucks

In Bezug auf die Reduzierung des Self-Justification-Drucks wurden in Abschnitt 2.3.3.1 auf Seite 35 drei Maßnahmen ermittelt. Diese Maßnahmen sind Tabelle 4.28 zu entnehmen.

Tabelle 4.28: Maßnahmen zur Reduzierung des Self-Justification-Drucks

Kurzbezeichnung	Formulierung	Wirkung
Mitverantwortung	Es gab andere Personen (z. B. Controller), die explizit Mitverantwortung für das Fehlschlagen des Projekts übernommen haben.	n. s.
Prozessqualität	Ich wurde nicht nur anhand der tatsächlichen Folgen meiner Entscheidungen bewertet, sondern auch für die Qualität meiner Entscheidungsprozesse (d. h., es wurde auch beurteilt, ob die Entscheidung zum jeweiligen Zeitpunkt gut begründet war).	n. s.
Ansehen	Mir wurde klar kommuniziert, dass mein Ansehen bei einem Projektabbruch keinerlei Schaden nehmen würde.	-0,20****

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

Von den drei untersuchten Maßnahmen wirkt nur der Schutz des Ansehens signifikant reduzierend auf den Self-Justification-Druck, dies jedoch mit einem vergleichsweise großen Pfadkoeffizienten von 0,20. Von allen hier diskutierten Wirkungen der Maßnahmen handelt es sich hierbei um die stärkste Wirkung einer Maßnahme. H_{M11} wird somit gestützt.²⁹⁰ Die anderen beiden Maßnahmen zeigen keine signifikante Wirkung, so dass H_{M9} und H_{M10} verworfen werden.

²⁸⁹ Vgl. Abschnitt 2.3.2.2 auf Seite 34 und Abbildung 2.3 auf Seite 28.

²⁹⁰ Vgl. Abschnitt 2.3.3.1 auf Seite 35 und Abbildung 2.3 auf Seite 28.

4.5.1.4 Maßnahmen zur Reduzierung des Sunk-Cost-Effekts

Zur Verringerung des Sunk-Cost-Effekts wurden in Abschnitt 2.3.4.1 auf Seite 38 fünf Maßnahmen hergeleitet, die in Tabelle 4.29 aufgeführt sind.

Tabelle 4.29: Maßnahmen zur Reduzierung des Sunk-Cost-Effekts

Kurzbezeichnung	Formulierung	Wirkung
expliziter Ausweis	Die Sunk Costs (d. h. in der Vergangenheit investierte Kosten, die nicht zurückgewonnen werden können) wurden ausdrücklich als solche ausgewiesen.	+0,09***
Irrelevanz gekennzeichnet	Die Sunk Costs wurden als irrelevant für Entscheidungen gekennzeichnet (d. h., es war deutlich, dass nur zukünftige Kosten und Erlöse entscheidungsrelevant sind).	n. s.
Verfahren	Mir standen angemessene Verfahren für die Bewertung des Projekts zur Verfügung.	n. s.
Expertenunterstützung	Ich wurde bei der Anwendung der Bewertungsinstrumente durch einen Experten (z. B. Controller) unterstützt.	n. s.
Alternativen	In der Projektberichterstattung wurde auf die Möglichkeit hingewiesen, die Ressourcen für das scheiternde Projekt in andere Projekte zu investieren.	+0,07*

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

Der explizite Ausweis der Sunk Costs und das Aufzeigen von Alternativen weisen signifikante Wirkungen auf. Beide Wirkungen sind jedoch im Gegensatz zu den Hypothesen positiv. Der explizite Ausweis der Sunk Costs steigert also den Sunk-Cost-Effekt, anstatt ihn zu verringern. Eine Erklärungsmöglichkeit wäre, dass dieser Effekt auftritt, wenn die Sunk Costs ausgewiesen werden, dem Entscheidungsverantwortlichen jedoch die Irrelevanz für die Entscheidungen nicht bewusst ist. Der hier nachgewiesene Effekt steht im Einklang mit der Literatur, die davor warnt, dass einzelne Maßnahmen durchaus unerwünschte Wirkungen haben können und oft erst das Zusammenspiel verschiedener Maßnahmen die kognitiven Biases wirkungsvoll bekämpfen kann.²⁹¹

Der Hinweis auf Alternativen im Projektbericht scheint ebenfalls den Sunk-Cost-Effekt zu verstärken, anstatt ihn zu verringern (allerdings nur auf einem Signifikanzniveau von 0,1). Die Gründe dafür sind aus den vorliegenden Daten nicht ersichtlich und sollten in zukünftigen Studien genauer untersucht werden, wenn sich die Messung einer positiven (d. h. verstärkenden) Wirkung auf den Sunk-Cost-Effekt replizieren lässt.

²⁹¹ Vgl. Zayer 2007, S. 215 ff.

Da alle Maßnahmen zur Reduzierung des Sunk-Cost-Effekts entweder keine signifikante oder sogar eine verstärkende Wirkung zeigen, müssen die Hypothesen H_{M12} bis H_{M16} verworfen werden.

4.5.1.5 Maßnahmen zur Reduzierung des übermäßigen Optimismus

Gestützt auf die Literatur und frühere Experimente,²⁹² wurden in der empirischen Erhebung zwei Maßnahmen abgefragt, die zur Reduzierung des übermäßigen Optimismus beitragen sollen. Diese sind in Tabelle 4.30 dargestellt.

Tabelle 4.30: Maßnahmen zur Reduzierung des übermäßigen Optimismus

Kurzbezeichnung	Formulierung	Wirkung
Analogien	Schätzungen über den voraussichtlichen Projekterfolg wurden mit Hilfe von Analogien/Vergleichen zu anderen Projekten durchgeführt.	n. s.
Advocatus Diaboli	Es gab jemanden, der die Rolle des „Advocatus Diaboli“ übernommen hat (d. h., jemand hat bewusst pessimistisch argumentiert).	n. s.

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

Weder die Verwendung von Analogien bei der Schätzung des Projekterfolgs noch der Einsatz eines Advocatus Diaboli zeigen eine signifikante Wirkung auf den übermäßigen Optimismus. Demzufolge müssen H_{M17} und H_{M18} verworfen werden.

4.5.1.6 Maßnahmen zur Reduzierung des nicht hinterfragten Freiraums

Von den in der Literatur vorgeschlagenen Maßnahmen²⁹³ zur Reduzierung des nicht hinterfragten Freiraums wurden drei Maßnahmen untersucht. Diese sind in Tabelle 4.31 auf der nächsten Seite angegeben.

Von den drei Maßnahmen erweisen sich die Bekanntgabe der Ziele und die frühzeitige Festlegung von Abbruchkriterien als signifikant wirksam. H_{M19} und H_{M21} werden somit durch die empirischen Ergebnisse gestützt. Die Diskussion der Ziele erweist sich als nicht signifikant, so dass H_{M20} verworfen wird.

²⁹² Vgl. Abschnitt 2.3.4.2 auf Seite 40.

²⁹³ Vgl. Abschnitt 2.3.5.2 auf Seite 42.

Tabelle 4.31: Maßnahmen zur Reduzierung des nicht hinterfragten Freiraums

Kurzbezeichnung	Formulierung	Wirkung
Zielbekanntgabe	Die Ziele, die ich mit der Weiterführung des Projekts verfolgt habe, wurden bekanntgegeben.	-0,12***
Zieldiskussion	Meine Projektziele wurden diskutiert (z. B. in Bezug auf ihre Erreichbarkeit).	n. s.
Abbruchkriterien	Bereits zu Beginn des Projekts (und z. B. an Meilensteinen) wurden Abbruchkriterien festgelegt, die den Projektbeteiligten bekannt waren (d. h. Zwischenziele, bei deren Verfehlen das Projekt abgebrochen werden sollte).	-0,13***

Signifikanzniveaus: * = 0,1; ** = 0,05; *** = 0,01; **** = 0,001; n. s. = nicht signifikant

4.5.1.7 Zusammenfassende Diskussion der Einzelwirkungen der Maßnahmen

Von den 21 Hypothesen²⁹⁴ über die Wirkung der verhaltensorientierten Maßnahmen konnten fünf Hypothesen bestätigt werden. 14 Maßnahmen zeigen keine signifikante Wirkung auf den ihnen jeweils zugeordneten kognitiven Faktor und zwei Maßnahmen zeigen eine signifikant verstärkende Wirkung des Faktors.

Der Mehrzahl der Maßnahmen kann somit keine allgemeingültige (d. h. kontextübergreifende) Wirkung bescheinigt werden. Die entgegengesetzte Schlussfolgerung, dass die Maßnahmen überhaupt keinen Nutzen zur Eskalationsreduzierung aufweisen, ist jedoch ebenfalls abzulehnen. Da die Wirksamkeit fast aller abgefragter Maßnahmen bereits in Experimenten gezeigt wurde,²⁹⁵ liegt eine andere Schlussfolgerung auf der Hand: Die Wirkung der Mehrzahl der Maßnahmen ist als kontextabhängig anzunehmen. Ob eine Maßnahme wirkt, hängt also von den ganz konkreten Bedingungen der jeweiligen Situation ab.²⁹⁶

Gleichzeitig ist es denkbar, dass die untersuchten Maßnahmen je nach Kontext auch für andere kognitive Faktoren als die hier jeweils untersuchten eine reduzierende Wirkung haben. Für diese Schlussfolgerung spricht die Tatsache, dass in der Literatur auch Quellen gefunden werden können, in denen einige Maßnahmen aus dem Modell von Zayer *anderen* kognitiven Faktoren

²⁹⁴ Vgl. Abbildung 2.3 auf Seite 28.

²⁹⁵ Vgl. Abschnitt 2.3 auf Seite 29.

²⁹⁶ Einige potentielle Kontextfaktoren sind im Fragebogen mit erhoben wurden; vgl. Abschnitt A.5 auf Seite 222. So wurden beispielsweise Fragen zur Projektkomplexität, dem Projektrisiko, der Abhängigkeit von äußeren Faktoren sowie zum Projektumfang (Manntage, Kosten und Kalenderzeit) gestellt. Für diese und weitere Aspekte wurde exploratorisch geprüft, ob sie die Wirkung der einzelnen Maßnahmen moderieren. Dabei wurde eine Reihe von moderierenden Einzeleffekten festgestellt. Eine klare Systematik war jedoch nicht erkennbar, weshalb auf eine detaillierte Diskussion der Ergebnisse an dieser Stelle verzichtet wird.

zugeordnet werden. So empfehlen beispielsweise Hammond/Keeney/Raiffa, zur Reduzierung des Sunk-Cost-Effekts die Entscheidungsqualität zu berücksichtigen;²⁹⁷ die gleiche Maßnahme wird von Zayer ausschließlich der Verringerung des Self-Justification-Effekts (und in der vorliegenden Arbeit dem Self-Justification-Druck) zugeordnet. Der Einsatz eines *Advocatus Diaboli* wird von Hammond/Keeney/Raiffa zur Reduzierung der selektiven Wahrnehmung vorgeschlagen, während er in Zayers Modell ausschließlich den übermäßigen Optimismus verhindern soll.

Die Schlussfolgerung der Kontextabhängigkeit der Maßnahmen gewinnt an Prägnanz, wenn sie den Analysen der kognitiven Faktoren gegenübergestellt wird. Wie in Abschnitt 4.4.1 gezeigt, ist die Wirkung der kognitiven Faktoren *kontextübergreifend* beobachtbar. Anders ausgedrückt: Den verhältnismäßig universalen, *kontextübergreifenden* kognitiven Ursachen der Eskalation des Commitments werden bislang überwiegend *kontextabhängige* Maßnahmen in der Literatur gegenübergestellt. Demzufolge besteht ein Bedarf an zukünftigen Untersuchungen, die genauer untersuchen, unter welchen Kontextbedingungen welche Maßnahme besonders geeignet ist.

Zu diesem Zweck bieten sich Fallstudien an, die dem Komplexitätsgrad der Zusammenhänge gerecht werden und insbesondere untersuchen können, an welchem Punkt eine bestimmte – auf theoretischer und experimenteller Basis empfohlene – Maßnahme in der Praxis ihre Wirksamkeit verlieren kann. Auch großzählige Befragungen können hilfreich sein, vor allem, wenn sie sich auf einige wenige Maßnahmen konzentrieren und diese mit einem hohen Detaillierungsgrad untersuchen.²⁹⁸

4.5.2 Wirkung aller Maßnahmen auf die Projekteskalation

Zayer weist darauf hin, dass erst die Verwendung eines umfassenden Maßnahmenkatalogs eine wirksame Reduzierung der kognitiven Faktoren ermöglicht.²⁹⁹ Aus diesem Grund wurde zusätzlich zur separaten Analyse jeder einzelnen Maßnahme ein Index (d. h. die Summe aller Maßnahmen) gebildet. Anschließend wurde die Wirkung dieses Index simultan auf alle kognitiven

²⁹⁷ Vgl. für diese und eine Reihe weiterer Maßnahmen Hammond/Keeney/Raiffa 2006, S. 122 f.

²⁹⁸ Es würde sich dabei anbieten, auch offene Fragen mit Freitextfeldern einzusetzen, um potentielle Probleme der jeweiligen Maßnahme exploratorisch zu erkunden.

²⁹⁹ Vgl. Zayer 2007, S. 218.

Faktoren und die Projekteskalation geschätzt. Das entsprechende Modell ist in Abbildung 4.12 dargestellt.

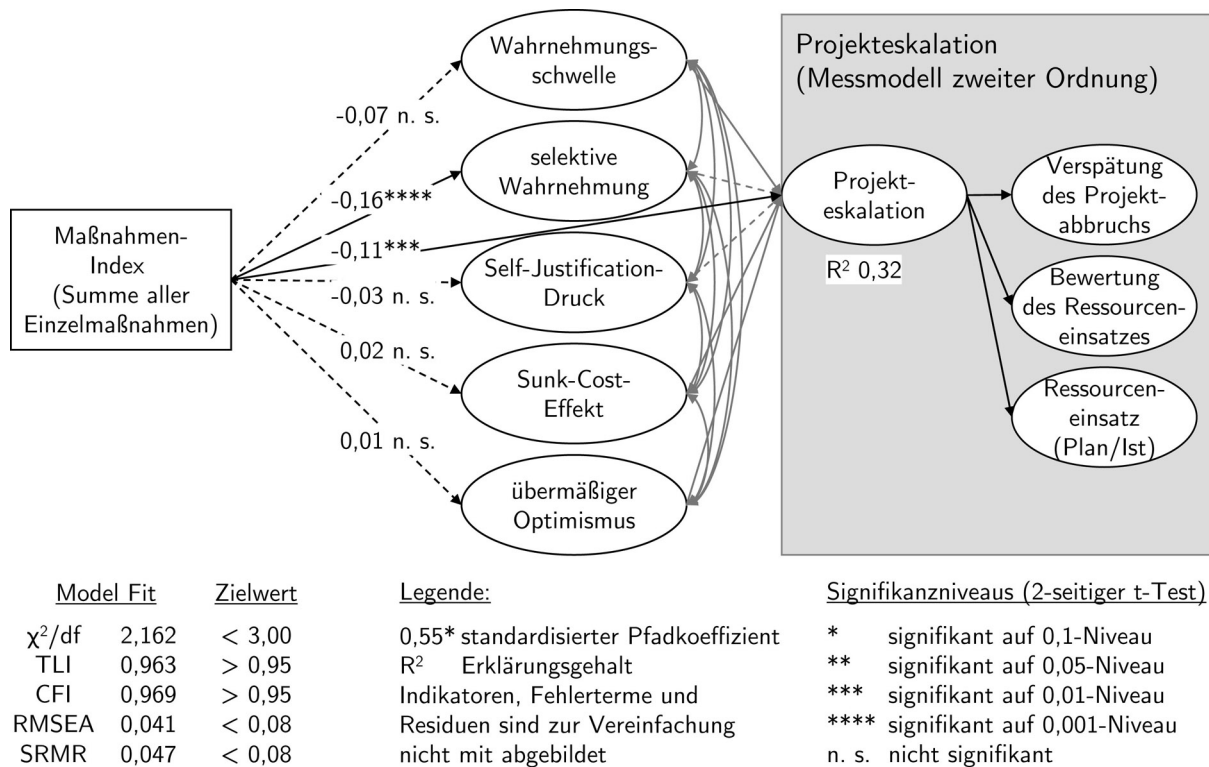


Abbildung 4.12: Wirkung aller Maßnahmen

Auf die vier Faktoren Wahrnehmungsschwelle, Self-Justification-Druck, Sunk-Cost-Effekt und übermäßiger Optimismus sind keine signifikanten Wirkungen durch die Gesamtmenge der Maßnahmen festzustellen. Eine Erklärung dafür könnte sein, dass die Wirkung jeder einzelnen Maßnahme kontextspezifisch ist und so im Mittel über alle Maßnahmen und Kontexte hinweg die Wirkungen nicht mehr als statistisch signifikant nachweisbar sind.

Der Index der Maßnahmen zeigt zwei signifikante Wirkungen. Zum einen kann die selektive Wahrnehmung gesenkt werden. Zum anderen wirkt der Index der Maßnahmen direkt signifikant reduzierend auf die Projekteskalation. Dieser direkte Effekt zeigt deutlich, dass die in der vorliegenden Studie untersuchten Wirkungsmechanismen der Maßnahmen nicht die einzigen relevanten (und möglicherweise auch nicht die wichtigsten) für die Eskalationsreduzierung sind. Die These aus Abschnitt 4.5.1.7 auf Seite 175, dass die Maßnahmen noch andere eskalationssenkende Wirkungen entfalten als auf den ihnen hier zugeordneten Faktor, wird damit gestützt.

Dadurch kann das bisher vorgestellte Bild weiter differenziert werden. Die Summe der untersuchten Maßnahmen trägt kontextübergreifend zur Reduzierung der Projekteskalation bei. Die

Theorien darüber, *wie* dies genau funktioniert, können durch die empirischen Daten jedoch nicht bestätigt werden. Anders ausgedrückt, die Literatur, auf der die vorliegende Studie basiert, erklärt die wesentlichen Wirkungsmechanismen zur Reduzierung der Eskalation noch nicht zufriedenstellend. Vielmehr scheinen die bisher gefundenen Erklärungen nur in bestimmten Kontexten zu greifen, während das Verständnis der kontextübergreifenden Mechanismen im Detail noch weitere Forschung erfordert.

4.5.3 Nutzung und Nutzen der Maßnahmen in der Unternehmenspraxis

Aus den bisherigen Analysen kann für die Praxis abgeleitet werden, dass fünf Maßnahmen eine signifikante Wirkung auf ihren jeweiligen kognitiven Faktor aufweisen. Dabei handelt es sich um das Festlegen von Meilensteinen, die Quantifizierung der Projektinformationen, den Schutz des Ansehens des Verantwortlichen im Falle eines Abbruchs, die Bekanntgabe der Ziele und das Ex-ante-Festlegen von Abbruchkriterien.³⁰⁰ Aus der signifikanten Wirkung dieser Maßnahmen kann keine Schlussfolgerung für ein einzelnes konkretes Unternehmen gezogen werden. Aber es liegt nahe, den Einsatz dieser Maßnahmen zumindest zu erwägen, da sie sich im Durchschnitt als wirksam erwiesen haben.

Wie die in Abbildung 4.12 auf der vorherigen Seite dargestellten Analysen andeuten, liegt der Wert der diskutierten Maßnahmen nicht allein im Reduzieren der kognitiven Faktoren, die im Fokus dieser Studie stehen. Das offene Bekanntgeben der Ziele kann beispielsweise auch zu mehr Akzeptanz des Projekts bei den Beteiligten führen. Da nicht alle relevanten Faktoren erfassbar sind, muss ein anderer Weg beschritten werden, um die Praxisrelevanz der Maßnahmen von einer übergeordneten Perspektive aus beurteilen zu können. Ein solcher Ansatzpunkt könnte die Überlegung sein, dass besonders erfolgreiche Unternehmen die einzelnen Maßnahmen in angemessenem Umfang einsetzen. Dies stellt selbstverständlich nur eine Annahme dar. Es könnte auch durchaus vorkommen, dass alle betrachteten Unternehmen die gleichen Fehler machen, oder auch, dass die besten Unternehmen manche Maßnahmen ungünstiger einsetzen, als Wettbewerber, die insgesamt bei der Projektsteuerung schlechter abschneiden. Darüber hinaus ist davon auszugehen, dass die optimale Einsatzhäufigkeit jeder Maßnahme von der spezifischen Situation des einzelnen Unternehmens beeinflusst wird.

³⁰⁰ Vgl. Abschnitt 4.5 auf Seite 169 ff.

Trotz aller Einschränkungen hinsichtlich der zulässigen Interpretationen, erscheint es grundsätzlich interessant, die Häufigkeit der durchgeführten Maßnahmen in der Praxis zu analysieren. Zu diesem Zweck werden im Folgenden jeweils die am besten gesteuerten zehn Prozent der Projekte mit den übrigen verglichen.³⁰¹ Als besonders erfolgreich im Sinne der Projektsteuerung werden hier die Unternehmen gewertet, die den Ressourcen-Einsatz hinsichtlich Manntagen, Kosten und Kalenderzeit jeweils „genau richtig“ bewertet haben.³⁰² Dies traf auf 68 der 677 Projekte, also (zufällig) genau zehn Prozent, zu. Abbildung 4.13 auf der nächsten Seite stellt die Rangfolge der Maßnahmen im Vergleich der beiden Projektgruppen dar.

Von den 21 untersuchten Maßnahmen zur Verbesserung der Projektsteuerung sind das Bekanntgeben der Ziele, das Abbilden des Projektverlaufs in konkreten Zahlen und das Festlegen von Meilensteinen die drei am häufigsten genutzten Maßnahmen. Diese gehören gleichzeitig zu den fünf Maßnahmen, deren kontextübergreifende Wirkung in den vorangegangenen Analysen signifikant war. Die Unternehmen setzen im Durchschnitt also die Maßnahmen am häufigsten ein, die auch die größte Wirksamkeit aufweisen. Dies deckt sich mit der Hypothese des Rationalitätsprinzips, dass Agenten angemessen (rational) auf ihre Problemsituation reagieren.³⁰³

Neben der Rangfolge der Einsatzhäufigkeiten können auch aus dem Vergleich zwischen erfolgreicher gesteuerten Projekten und weniger erfolgreich gesteuerten Projekten Rückschlüsse gezogen werden. Besonders starke Unterschiede zwischen den beiden Gruppen treten im Hinblick darauf auf, dass bei den erfolgreicher gesteuerten Projekten häufiger eindeutige Kriterien zur Beurteilung des Projekterfolgs verwendet werden, vermehrt die Prozessqualität der Entscheidungen beurteilt und stärker das Ansehen des Projektverantwortlichen geschützt wird. Von diesen drei Maßnahmen hatte sich der Schutz des Ansehens als kontextübergreifend wirksam gezeigt.

Abschließend stellt sich die Frage, ob der Nutzen der Maßnahmen auch *quantifiziert* werden kann, etwa näherungsweise über den relativen Ressourcen-Einsatz. Die erfolgreichsten zehn Prozent der Projekte unterscheiden sich signifikant von den restlichen Projekten im Hinblick auf den relativen Ressourcen-Einsatz.³⁰⁴ Eine Größenordnung der Unterschiede vermittelt Abbildung 4.14. So liegen die durchschnittlich investierten Manntage bei den erfolgreicher gesteuerten Projekten bei 58 Prozent, bei den restlichen hingegen bei 94 Prozent.

³⁰¹ Die Mittelwerte und Standardabweichungen für das *gesamte* Sample sind in Tabelle A.3 auf Seite 239 im Anhang aufgeführt.

³⁰² Vgl. Abschnitt 4.3.1.2 auf Seite 125.

³⁰³ Vgl. Koertge 1975, S. 440, Koertge 1979, S. 87, Popper 1985, S. 357 ff., und Popper 1994, S. 169.

³⁰⁴ Vgl. Abschnitt 4.3.1.3 auf Seite 126.

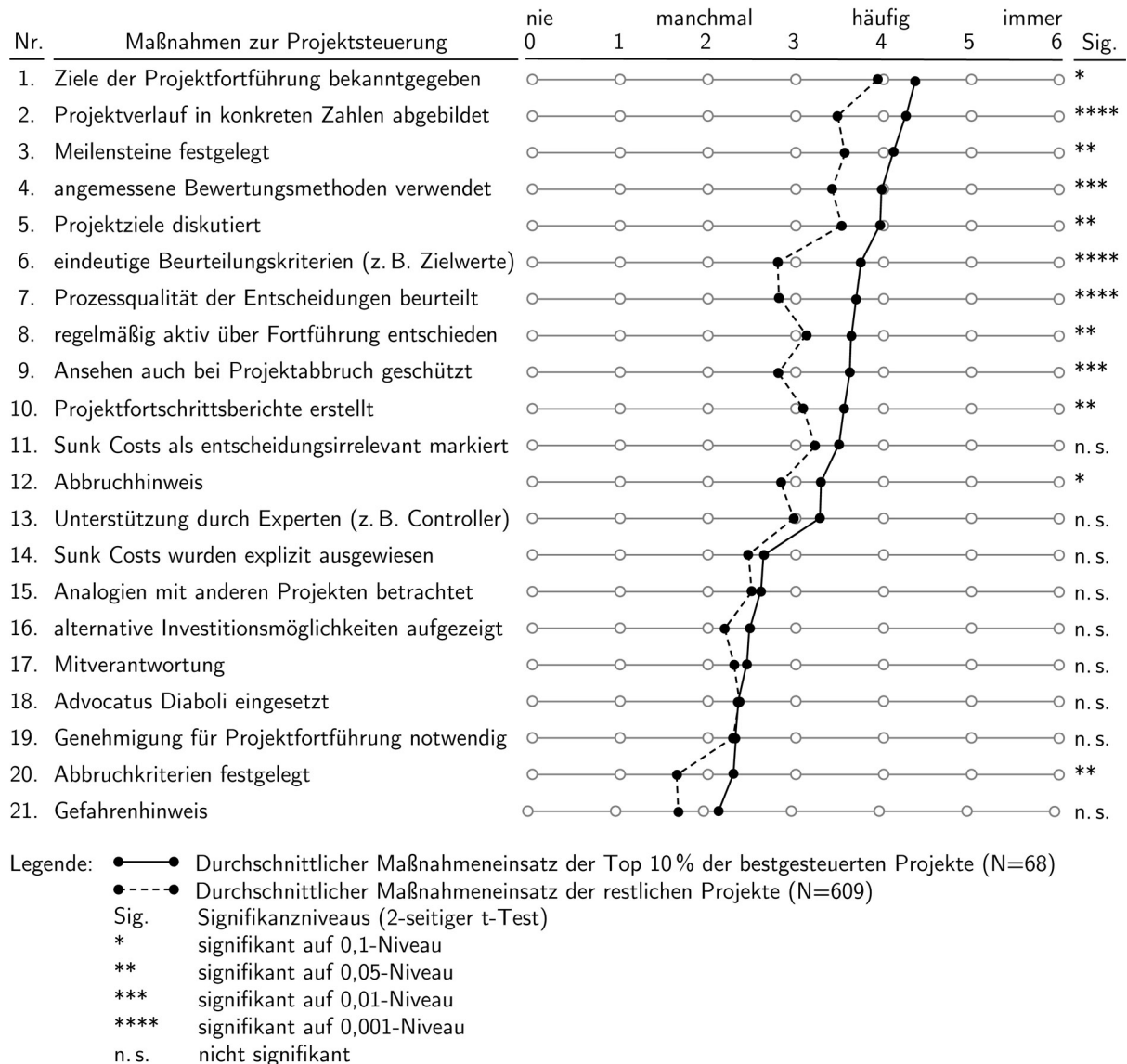


Abbildung 4.13: Vergleich der Häufigkeit der Maßnahmenverwendung

Hinsichtlich der bis zum Abbruch investierten Kosten liegen die besser gesteuerten Projekte durchschnittlich bei 51 Prozent im Vergleich zur geplanten Gesamtinvestitionssumme, die restlichen Projekte bei 88 Prozent. Für die besser gesteuerten Projekte liegt die aufgewendetet Investitionssumme also durchschnittlich 37 Prozent niedriger. Dabei umfasst die Angabe der restlichen Projekte sowohl die Projekte, für die zu geringe Investitionen diagnostiziert wurden, als auch die mit zu hohen Investitionen. Vergleicht man die Top zehn Prozent der Projekte nur mit den Projekten, bei denen die Befragten einen zu hohen Ressourceneinsatz konstatiert haben, ist

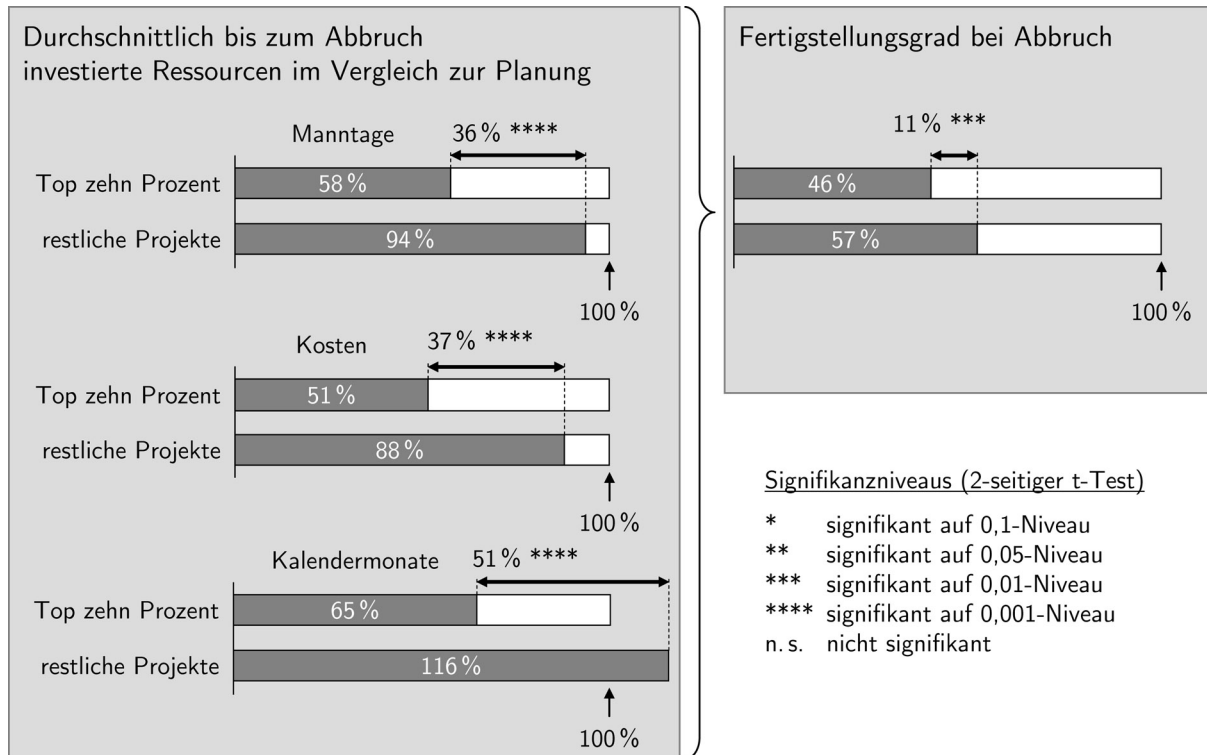


Abbildung 4.14: Investierte Ressourcen und Fertigstellungsgrad

der Unterschied noch größer, nämlich etwa 48 Prozent. Bei einem durchschnittlich geplanten Projektvolumen von 68 Mio. Euro³⁰⁵ entspricht dies einer Differenz von knapp 33 Mio. Euro.

In Bezug auf die bis zum Abbruch verstrichene Kalenderzeit liegen die besser gesteuerten Projekte bei 65 Prozent, die restlichen Projekte bei 116 Prozent. Das bedeutet, dass ein Großteil der Projekte erst dann abgebrochen wird, wenn die ursprünglich geplante Zeit bis zur Fertigstellung bereits um 16 Prozent überschritten worden ist.

Der Fertigstellungsgrad liegt bei den besser gesteuerten Projekten bei 46 Prozent, bei den restlichen bei 57 Prozent. Die Differenz hinsichtlich des Fertigstellungsgrades in Höhe von 11 Prozent ist dabei deutlich geringer als die Differenzen hinsichtlich der eingesetzten Ressourcen. Die besser gesteuerten Projekte weisen also ein günstigeres Verhältnis von Aufwand zu Fertigstellung auf (waren also effizienter). Gleichzeitig wurden sie in Bezug auf den Fertigstellungsgrad zu einem früheren Zeitpunkt abgebrochen.

³⁰⁵ Vgl. Abschnitt 4.1.3 auf Seite 107.

4.6 Prüfung potentieller Fehlerquellen

4.6.1 Prüfung des Informant-Bias

Um den Informant-Bias prüfen zu können, wurde in der Befragung nach der Entscheidungskompetenz bezüglich Abbruch oder Fortsetzung des Projekts gefragt. Die Befragten können daher in zwei Gruppen eingeteilt werden. Gruppe 1 (N = 287) umfasst die Entscheidungsverantwortlichen und Gruppe 2 (N = 390) die Nicht-Entscheidungsverantwortlichen. Die Ergebnisse des Gruppenvergleichs sind in Abbildung 4.15 dargestellt.

Das Modell besitzt 435 Freiheitsgrade und einen χ^2 -Wert von 682,476. Alle Gütemaße erfüllen die geforderten Grenzwerte. Die Ergebnisse des Gruppenvergleichs zeigen drei wesentliche Unterschiede hinsichtlich der Wirkungen der kognitiven Faktoren. (1) In der Gruppe der Entscheidungsverantwortlichen ist die Wirkung der Wahrnehmungsschwelle signifikant, bei den Nicht-Entscheidungsverantwortlichen hingegen nicht. (2) Der Sunk-Cost-Effekt zeigt bei den Nicht-Entscheidungsverantwortlichen eine mehr als doppelt so starke Wirkung auf die Projekteskalation als bei den Entscheidungsverantwortlichen. (3) Für die Gruppe der Entscheidungsverantwortlichen zeigt der übermäßige Optimismus keine signifikante Wirkung, bei den Nicht-Entscheidungsverantwortlichen hingegen eine starke Wirkung mit einem standardisierten Pfadkoeffizienten von 0,32, der auf 0,01-Niveau signifikant ist.

(1) Hinsichtlich der Wirkung der Wahrnehmungsschwelle führt ein χ^2 -Differenztest zu dem Ergebniss, dass der Unterschied zwischen den beiden Gruppen *nicht* signifikant ist.³⁰⁶ Dessen ungeachtet ist es theoretisch denkbar, dass sich die Messung zwischen beiden Gruppen unterscheidet. Ein möglicher Grund wäre beispielsweise, dass Menschen dazu neigen, Erfolg sich selbst zuzuschreiben, Misserfolg jedoch auf externe Umstände zurückzuführen.³⁰⁷ Eine Betrachtung der Konstrukte für die kognitiven Faktoren zeigt, dass die Indikatoren zur Messung Wahrnehmungsschwelle (vgl. Tabelle 4.9 auf Seite 131) möglicherweise als extern (mit)verursacht angesehen könnten, während dies bei den anderen Konstrukten weniger der Fall ist. Demzufolge würden Entscheidungsträger tendentiell die Wahrnehmungsschwelle um so stärker bewerten, je stärker sie die Projekteskalation einschätzen. Diese Hypothese könnte in Folgestudien mit Hilfe des Konstrukts „Locus of Control“ geprüft werden.³⁰⁸ Dieses Konstrukt erfasst, ob der

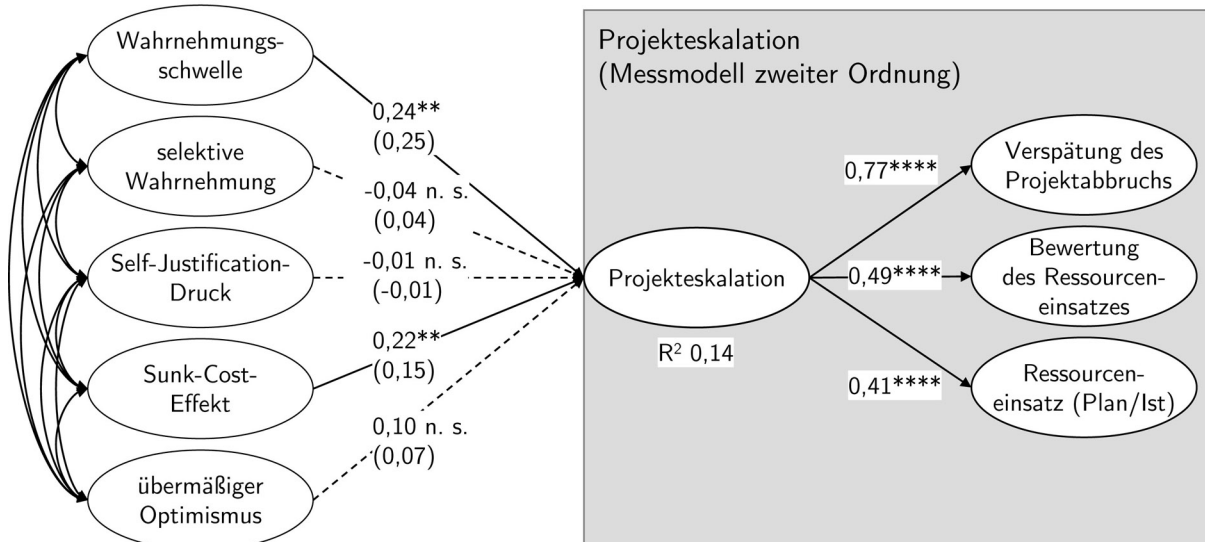
³⁰⁶ Für eine Erläuterung des χ^2 -Differenztests für die Untersuchung von moderierenden Effekten vgl. Abschnitt 4.4.2 auf Seite 162.

³⁰⁷ Vgl. beispielsweise Gervais/Odean 2001, S. 1.

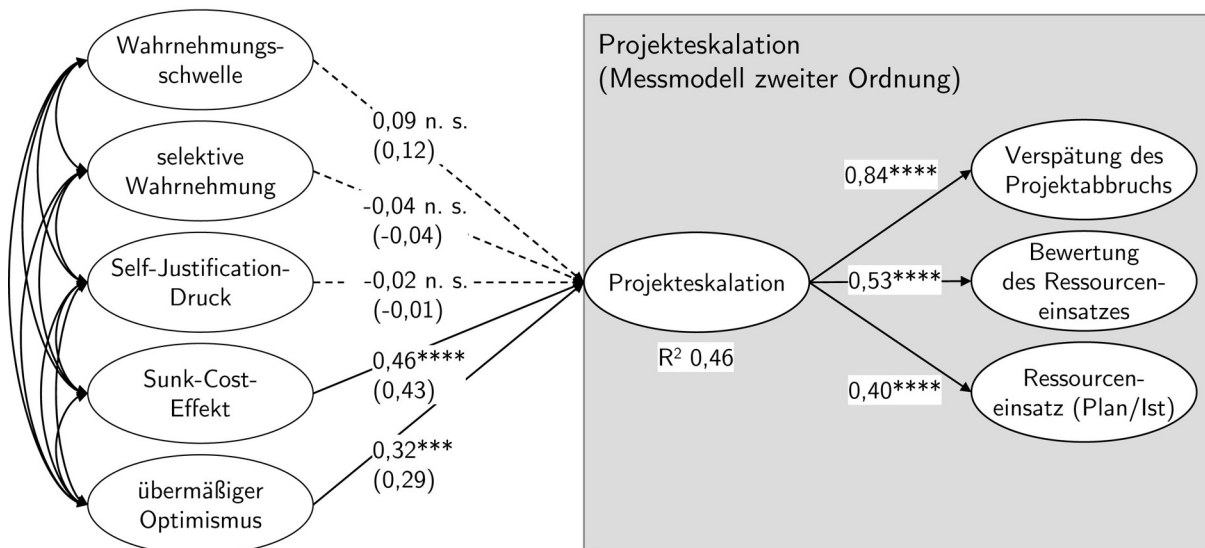
³⁰⁸ Vgl. für das Konstrukt „Locus of Control“ Bruner/Hensel 1992, S. 346 f. und die dort angegebene Literatur.

Respondent die ihm widerfahrenden Ereignisse eher sich selbst oder eher externen Ursachen zuschreibt. Würde dieses Konstrukt mit der Wahrnehmungsschwelle korrelieren, würde dies die geäußerte Hypothese stützen.

Gruppe 1 (N = 287): Entscheidungsverantwortliche



Gruppe 2 (N = 390): Nicht-Entscheidungsverantwortliche



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)	
χ^2/df	1,569	< 3,00	0,55* standardisierter Pfadkoeffizient	* signifikant auf 0,1-Niveau
TLI	0,967	> 0,95	(0,55*) unstandardisierter Pfadkoeffizient	** signifikant auf 0,05-Niveau
CFI	0,972	> 0,95	R ² Erklärungsgehalt	*** signifikant auf 0,01-Niveau
RMSEA	0,029	< 0,08	Indikatoren, Fehlerterme und Residuen sind zur Vereinfachung nicht mit abgebildet	**** signifikant auf 0,001-Niveau
SRMR	0,059	< 0,08		n. s. nicht signifikant

Abbildung 4.15: Gruppenvergleich Entscheidungsverantwortliche vs. Nicht-Entscheidungsverantwortliche für Modell 2

Neben der fehlerhaften Kausalattribution könnte auch das Bedürfnis, das Gesicht zu wahren, die Messung beeinflussen. Stimmt die Hypothese, dass die Wahrnehmungsschwelle stärker als andere Faktoren als extern angesehen wird, so läge es nahe, ein schlechtes Ergebnis (hier die Projekteskalation) mit einer hohen Wahrnehmungsschwelle zu rechtfertigen. Die statistischen Untersuchungen der vorliegenden Daten zeigen allerdings, dass die Wahrnehmungsschwelle negativ mit dem Kontrollkonstrukt Impression-Management korreliert.³⁰⁹ Dies spricht gegen diesen zweiten Erklärungsansatz.

(2) Der Unterschied in der Wirkungsstärke des Sunk-Cost-Effekts ist zwischen den beiden Gruppen signifikant.³¹⁰ Unterschiedlich ist allerdings lediglich die Stärke des Effekts – die Richtung ist bei beiden Gruppen gleich. Die Beurteilung des Sunk-Cost-Effekts als wichtiger Einflussfaktor für die Projekteskalation ändert sich durch die Gruppenunterschiede nicht.

Eine mögliche Erklärung für die geringe Wirkungsstärke bei der Gruppe der Entscheidungsverantwortlichen liegt darin, dass eine Person, die einem kognitiven Bias unterliegt, dies möglicherweise weniger gut beurteilen kann als außenstehende Personen. Ist die Messung des kognitiven Bias weniger präzise, kann der geschätzte Pfadkoeffizient eine geringere Wirkungsstärke zeigen, als es der Realität entsprechen würde. Dies würde auch die Unterschiede hinsichtlich der erklärten Varianz (R^2) plausibel erscheinen lassen: Für die Gruppe der Nicht-Verantwortlichen ergibt sich ein dreimal höheres R^2 als bei den Entscheidungsverantwortlichen.

(3) Auch für den signifikanten Unterschied hinsichtlich des übermäßigen Optimismus³¹¹ stellt die unter Punkt (2) angeführte Argumentation eine mögliche Erklärung dar. Andererseits ist es allerdings auch denkbar, und dies gilt auch für alle anderen Variablen, dass die Nicht-Entscheidungsverantwortlichen weniger Einblick in die tatsächliche Lage des Entscheidungsverantwortlichen haben und somit ungenauere Angaben machen als der Entscheidungsträger. Welcher der beiden Respondentengruppen die objektiveren Angaben zuzuschreiben sind, ist auf der Basis der vorliegenden Daten nicht zu klären.

Eine andere potentielle Quelle für Unterschiede zwischen den Gruppen besteht darin, dass die Gruppen sich möglicherweise systematisch im Hinblick auf die Projekte unterscheiden, über die sie Auskunft geben. So wäre es beispielsweise denkbar, dass die Entscheidungsverantwortlichen in höheren Positionen sind und sich deshalb im Durchschnitt mit größeren Projekten beschäftigen, als die Nicht-Entscheidungsverantwortlichen. Um dies zu prüfen, wurden die Mit-

³⁰⁹ Vgl. Abschnitt 4.6.2.2 auf Seite 187.

³¹⁰ Für die Wirkung des Sunk-Cost-Effekts ergibt der χ^2 -Differenztest: $\chi^2_{(1)} = 6,880$; $p = 0,009$.

³¹¹ Für die Wirkung des übermäßigen Optimismus ergibt sich $\chi^2_{(1)} = 7,672$; $p = 0,006$.

telwerte für die Projektgröße (Manntage, Kosten und Kalenderzeit) beider Gruppen verglichen. Die durchgeführten t-Tests ergaben jedoch keine signifikanten Unterschiede.

Abgesehen von den unter (2) und (3) diskutierten Zusammenhängen, unterscheiden sich alle übrigen Pfadkoeffizienten des Modells in Abbildung 4.15 auf Seite 183 *nicht* signifikant zwischen den beiden Gruppen. Zusätzlich wurde auch für ein Modell einschließlich des Index aller Maßnahmen ein Gruppenvergleich durchgeführt,³¹² um mögliche Unterschiede hinsichtlich der Wirkung der Maßnahmen zu prüfen. Weder die Wirkung des Maßnahmen-Index auf die kognitiven Faktoren noch auf die Projekteskalation unterscheiden sich signifikant zwischen Entscheidungsverantwortlichen und Nicht-Entscheidungsverantwortlichen.

Insgesamt zeigen die Ergebnisse, dass die Berücksichtigung des Informant-Bias die vorherigen Analysen zur Wirkung der kognitiven Faktoren nicht grundlegend in Frage stellt. Jedoch bestätigen die Unterschiede hinsichtlich der Wirkung des übermäßigen Optimismus, dass die Befragung von nur einer Art von Informanten zu einseitigen Ergebnissen führen kann. Deshalb ist es auch für nachfolgende Studien zur Projekteskalation empfehlenswert, ein Multi-Informant-Design zu verwenden, wie dies in der vorliegenden Studie der Fall war. Noch klarer könnten die Einflüsse des Informant-Bias in zukünftigen Studien herausgearbeitet werden, wenn eine dyadische Befragung zur Anwendung käme, also mehr als eine Person zum jeweils gleichen Projekt befragt wird.

4.6.2 Prüfung der Common-Method-Biases

4.6.2.1 Prüfung der Methodenvarianz mit der Marker-Variablen

In einer Meta-Analyse von 70 empirischen Studien fanden Cote/Buckley heraus, dass die Methodenvarianz je nach Fachgebiet und Art des Konstrukts zwischen 15 und 41 Prozent lag.³¹³ Es erscheint deshalb unbedingt angebracht, die empirischen Daten auf potentielle Verzerrungen durch einen Methodenbias zu prüfen. In der Literatur werden dazu vor allem vier Techniken vorgeschlagen: (1) die traditionelle Multitrait-Multimethod-Prozedur (MTMM), (2) die moderne MTMM-Variante unter Einsatz einer konfirmatorischen Faktorenanalyse, (3) Harmans Single-Factor-Test und (4) die Verwendung von Marker-Variablen.³¹⁴ Die MTMM-Technik auf

³¹² Vgl. Abbildung 4.12 auf Seite 177.

³¹³ Vgl. Cote/Buckley 1987.

³¹⁴ Vgl. Malhotra/Kim/Patil 2006, S. 1866 ff.

der Basis der konfirmatorischen Faktorenanalyse und die Marker-Variablen-Technik gelten als die zuverlässigsten State-of-the-Art-Methoden.³¹⁵ Beide erfordern jedoch, dass das Forschungsdesign schon vor der Erhebung der Daten die spätere Methodenprüfung berücksichtigt. Dies war bei der vorliegenden Studie der Fall.³¹⁶ Im Folgenden wird die Methodenprüfung mit der Marker-Variablen-Technik durchgeführt.

Die folgenden Berechnungen stützen sich auf die Arbeiten von Lindell/Whitney und Malhotra/Kim/Patil.³¹⁷ Zunächst wurde eine konfirmatorische Faktorenanalyse durchgeführt, die alle Messmodelle einschließlich des Method-Variance-Markers (MV-Markers) „Kapitalmarktorientierung“ umfasste.³¹⁸ Die Ergebnisse zeigen, dass das Modell die Daten gut repräsentiert:³¹⁹ $\chi^2_{(838)}$: 1522,01, TLI: 0,96, CFI: 0,96, RMSEA: 0,035 und SRMR: 0,034.

Ziel des MV-Marker-Verfahrens ist es, die statistischen Zusammenhänge zwischen den Modellvariablen um die Methodenvarianz zu bereinigen. Die bereinigten Korrelationen lassen sich nach der folgenden Formel ermitteln:³²⁰

$$r_A = \frac{r_U - r_M}{1 - r_M} \quad (4.3)$$

Dabei ist r_A die CMV-adjustierte Korrelation der Modellvariablen, r_U die nicht adjustierte (uncorrected) Korrelation der Modellvariablen und r_M der Anteil der Korrelation, der durch die Methodenvarianz verursacht wird.

Um die Methodenvarianz zu bestimmen, wurden die Korrelationen des MV-Markers mit allen anderen Konstrukten betrachtet. Mit der Mehrzahl der Konstrukte weist die Kapitalmarktorientierung nur eine sehr geringe, nicht signifikante Korrelation auf. Als Indikator der Methodenvarianz kann die *kleinste* Korrelation zwischen zwei Variablen herangezogen werden.³²¹ Für das vorliegende Sample ist die Korrelation zwischen Kapitalmarktorientierung und dem Sunk-Cost-Effekt kleiner als 0,0005. Diese Korrelation ist so gering, dass die Methodenvarianz für die Un-

³¹⁵ Vgl. Malhotra/Kim/Patil 2006, S. 1867 f.

³¹⁶ Vgl. Abschnitt 3.3.4 auf Seite 66.

³¹⁷ Vgl. Lindell/Whitney 2001 und Malhotra/Kim/Patil 2006.

³¹⁸ Vgl. zur Methodenvarianz Abschnitt 3.3.4 auf Seite 66 und für eine Erläuterung des MV-Markers „Kapitalmarktorientierung“ Abschnitt 4.3.4.1 auf Seite 139.

³¹⁹ Für die Anspruchsniveaus der Gütekriterien vgl. Tabelle 4.4 auf Seite 120.

³²⁰ Vgl. Lindell/Whitney 2001, S. 116, und Malhotra/Kim/Patil 2006, S. 1868.

³²¹ Vgl. Lindell/Whitney 2001, S. 115.

tersuchung als irrelevant gelten kann. Von den weiteren Schritten der von Lindell/Whitney und Malhotra/Kim/Patil vorgeschlagenen Prozedur wird deshalb abgesehen.³²²

4.6.2.2 Prüfung der sozialen Erwünschtheit und des Impression-Managements

Für die Prüfung der sozialen Erwünschtheit wurden zwei Wege beschritten.³²³ (1) Zum einen wurde ein Index aus fünf Fragen zum Impression-Management gebildet.³²⁴ (2) Zum anderen wurde das Konstrukt zur selektiven Wahrnehmung sowohl direkt als auch indirekt abgefragt,³²⁵ was dem MTMM-Verfahren entspricht.

Impression-Management-Index

Zur Prüfung der potentiellen Wirkung des Impression-Managements wurde das Modell 2 für die Gruppe der Entscheidungsverantwortlichen einmal ohne (vgl. Abbildung 4.16 auf der nächsten Seite) und einmal mit dem Impression-Management-Index geschätzt (vgl. Abbildung 4.17 auf Seite 189).

Der Impression-Management-Index weist mit der Wahrnehmungsschwelle und der selektiven Wahrnehmung signifikante negative Korrelationen auf.³²⁶ Dies entspricht dem erwarteten Effekt: Je stärker die Tendenz einer Person, sich in erwünschter Weise zu präsentieren, desto schwächer wird sie ihre kognitiven Biases in einer Befragung bewerten. Die Korrelationen zwischen dem Impression-Management-Index und allen anderen kognitiven Faktoren sind nicht signifikant.

Aus dem Vergleich der beiden Modelle geht hervor, dass das Impression-Management keinen Einfluss auf das Eskalationsmodell hat. Der Impression-Management-Index weist keine signifikante Wirkung auf die Projekteskalation auf. Darüber hinaus bleiben alle Pfadkoeffizienten in ihrer Signifikanz und ihrer Wirkungsstärke gleich. Auch das R^2 der Projekteskalation verändert sich nicht durch das Berücksichtigen des Impression-Managements.

³²² Vgl. Lindell/Whitney 2001 und Malhotra/Kim/Patil 2006.

³²³ Vgl. hierzu Abschnitt 3.3.4.1 auf Seite 68.

³²⁴ Vgl. Tabelle 4.19 auf Seite 140.

³²⁵ Vgl. Tabelle 4.21 auf Seite 142.

³²⁶ Die Korrelation zwischen dem Impression-Management-Index und der Wahrnehmungsschwelle beträgt -0,11 (t-Wert = -1,753; $p = 0,08$). Die Korrelation mit der selektiven Wahrnehmung beträgt -0,18 (t-Wert = -2,676; $p = 0,007$).

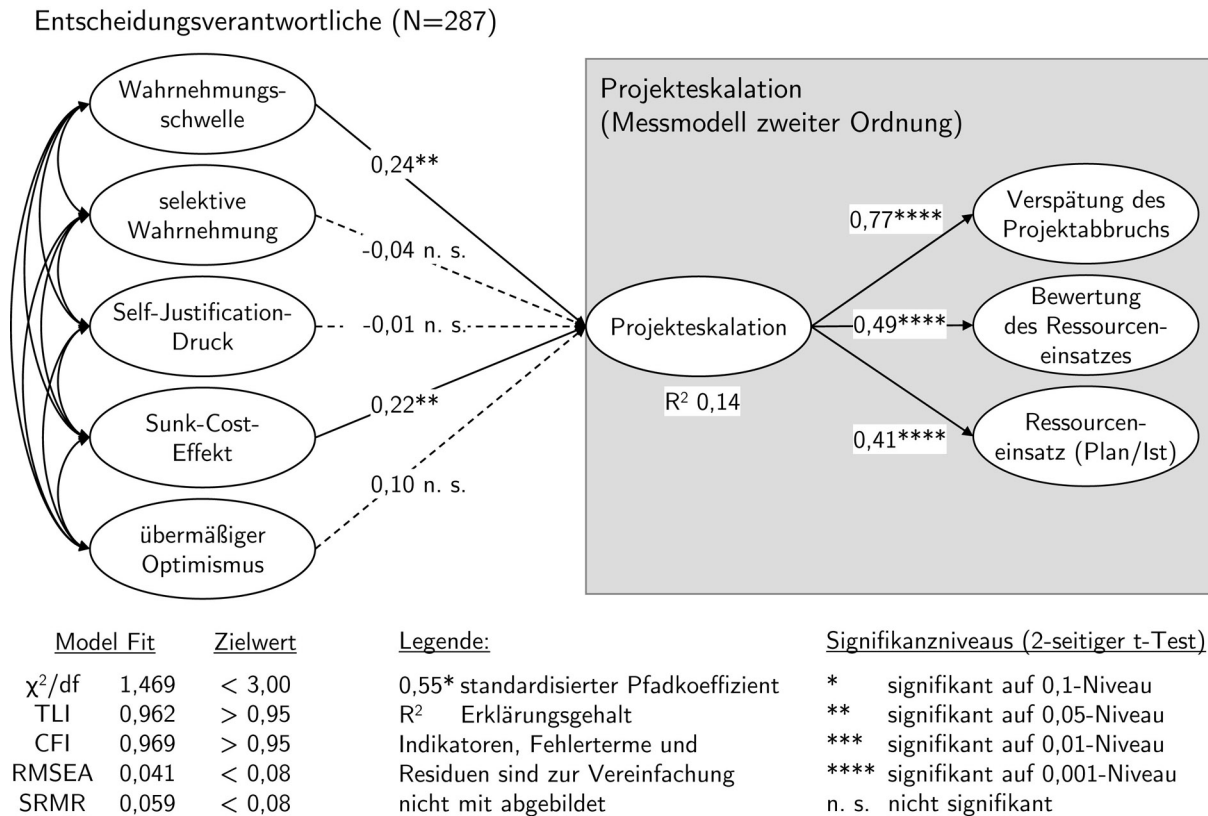


Abbildung 4.16: Eskalation (Modell 2) ohne Impression-Management-Index

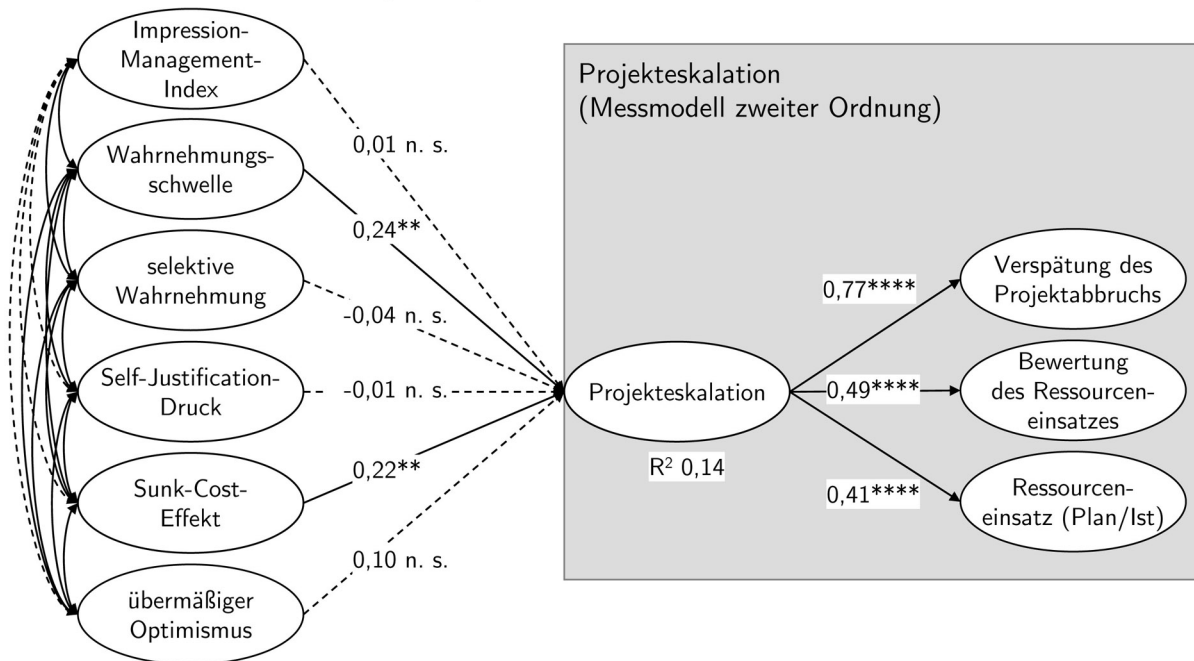
Multitrait-Multimethod-Verfahren

Für die vorliegende Untersuchung wurde das MTMM-Design für die Variable selektive Wahrnehmung eingesetzt.³²⁷ Um diesen Ansatz zu testen, wurde eine konfirmatorische Faktorenanalyse durchgeführt. Das Messmodell enthält sechs Indikatoren: drei Indikatoren der Selbstauskunft und drei Indikatoren der Auskünfte über einen anderen Manager. Das Messmodell wurde so spezifiziert, dass ein Faktor auf alle sechs Indikatoren lädt. Ein zweiter Faktor wurde so spezifiziert, dass er nur auf die drei Indikatoren der Selbstauskunft lädt. Bei einer solchen Spezifizierung erklärt (unter den genannten Annahmen) der erste Faktor die „tatsächliche Einschätzung der selektiven Wahrnehmung“. Der zweite Faktor (der Methodenfaktor) erklärt den Anteil der Varianz in der Selbstauskunft, die durch die soziale Erwünschtheit verursacht wird.

Die Analyse des Modells zeigt, dass es in der Tat einen gemeinsamen Faktor für beide Messungen gibt, beide Methoden also zum Teil dasselbe messen. Allerdings sind die Faktorladungen auf den Methodenfaktor größer als die Faktorladungen für den gemeinsamen Faktor. Folgt man den Annahmen von Jo, so spricht dies für die Existenz eines starken Social-Desirability-Bias.

³²⁷ Für eine ausführlichere Erläuterung vgl. Seite 71.

Entscheidungsverantwortliche (N=287)



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df	1,398	< 3,00	* signifikant auf 0,1-Niveau
TLI	0,965	> 0,95	** signifikant auf 0,05-Niveau
CFI	0,971	> 0,95	*** signifikant auf 0,01-Niveau
RMSEA	0,037	< 0,08	**** signifikant auf 0,001-Niveau
SRMR	0,057	< 0,08	n. s. nicht signifikant

Abbildung 4.17: Eskalation (Modell 2) mit Impression-Management-Index

In einem nächsten Schritt müsste beurteilt werden, inwieweit das Vorliegen eines solchen Methodeneffekts die Wirkungszusammenhänge zwischen unabhängiger und abhängiger Variablen im Strukturmodell beeinflusst. Üblicherweise wird damit geprüft, ob ein scheinbar signifikanter Zusammenhang nur auf Grund des Methodeneffekts als signifikant erscheint. Im vorliegenden Fall (vgl. Abbildung 4.5 auf Seite 150) liegt jedoch kein signifikanter Effekt für die selektive Wahrnehmung auf die Projekteskalation vor. Es wurde deshalb stattdessen geprüft, ob der Effekt signifikant wird, wenn die selektive Wahrnehmung gemäß der Methode von Jo im Gesamtmodell 2 spezifiziert wird. Dies trifft nicht zu.

Im vorliegenden Fall kann der Methodeneffekt also als unerheblich für die untersuchten Zusammenhänge betrachtet werden. Für die anderen kognitiven Faktoren wurde das MTMM-Verfahren nicht eingesetzt. Dies wäre für Folgeuntersuchungen zu empfehlen, wenn keine neuen Erkenntnisse dazu führen, die Methode von Jo grundsätzlich zu verwerfen.

4.6.2.3 Prüfung des Leniency-Bias

Unter dem Leniency-Bias wird die Tendenz verstanden, andere Personen in sozial erwünschter Weise zu beschreiben.³²⁸ Zur Kontrolle des Leniency-Bias waren drei Verfahren in der vorliegenden Studie berücksichtigt worden. Das Multitrait-Multimethod-Verfahren führt nicht zu klaren Ergebnissen, weil das dafür verwendete Konstrukt, die selektive Wahrnehmung, nicht signifikant auf die Projekteskalation wirkt.³²⁹ Im Folgenden werden deshalb die Ergebnisse der Prüfungen mit dem Leniency-Index und dem Freundschaftskonstrukt beschrieben.

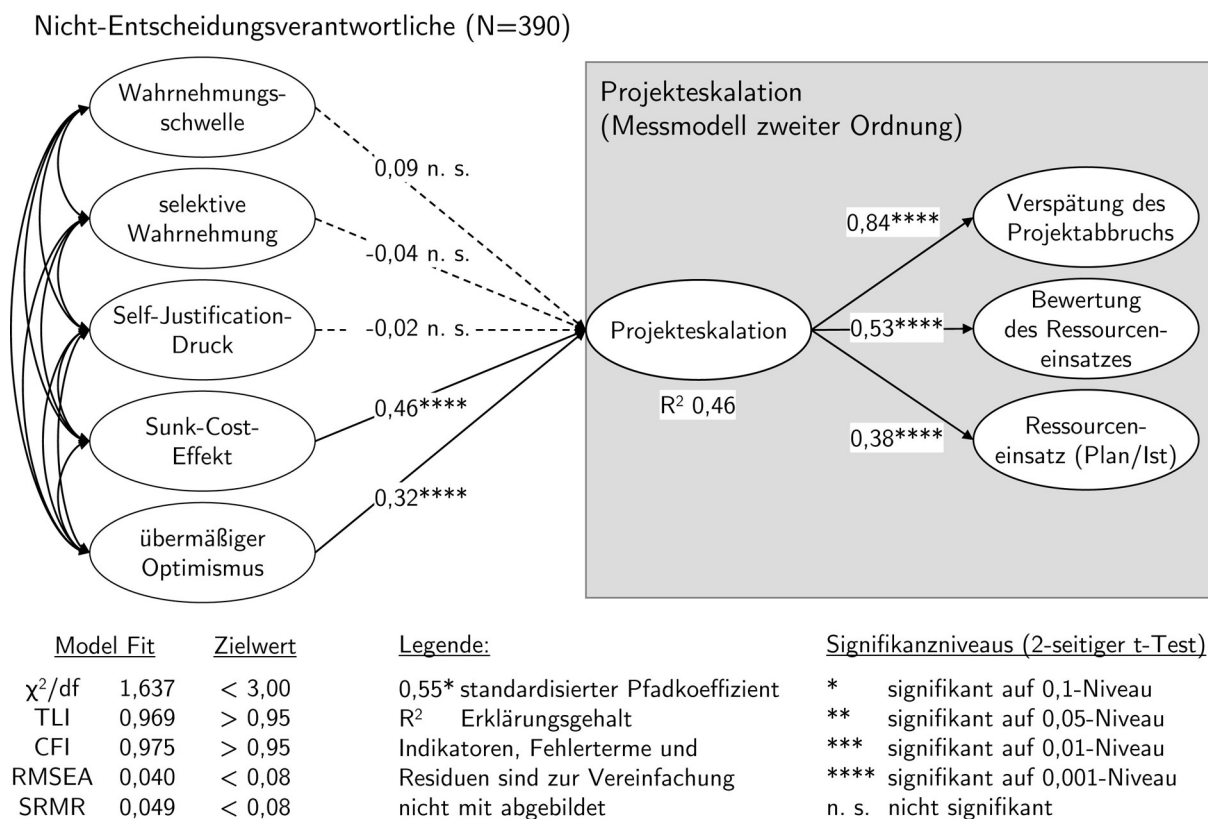


Abbildung 4.18: Eskalation (Modell 2), beurteilt durch Nicht-Entscheidungsverantwortliche

Beide Verfahren nutzen jeweils eine Kontrollvariable und prüfen, ob sich die Aussage des Gesamtmodells ändert, wenn die jeweilige Variable mit modelliert wird. Als Bezugspunkt für den Vergleich dient dabei jeweils das Eskalationsmodell 2 aus Abschnitt 4.4.1.2 auf Seite 150. Da der Leniency-Bias nur Auskünfte über andere Personen betrifft und keine Selbstauskünfte, fließen nur die Aussagen der Nicht-Entscheidungsverantwortlichen (N = 390) in die Analysen ein.

³²⁸ Vgl. ausführlicher hierzu und zu den weiteren Ausführungen Abschnitt 3.3.4.2 auf Seite 72.

³²⁹ Vgl. Seite 189.

Die geschätzten Parameter für das Ausgangsmodell sind in Abbildung 4.18 auf der vorherigen Seite dargestellt.

Leniency-Index

Die erste Kontrollvariable stellt der Leniency-Index dar. Der Leniency-Index ist analog zum Impression-Management-Index aufgebaut.³³⁰ Er umfasst die gleichen Fragen mit dem Unterschied, dass statt einer Selbstauskunft stets um die Auskunft über den Entscheidungsverantwortlichen gebeten wird. Das Modell einschließlich Leniency-Index ist in Abbildung 4.19 dargestellt.

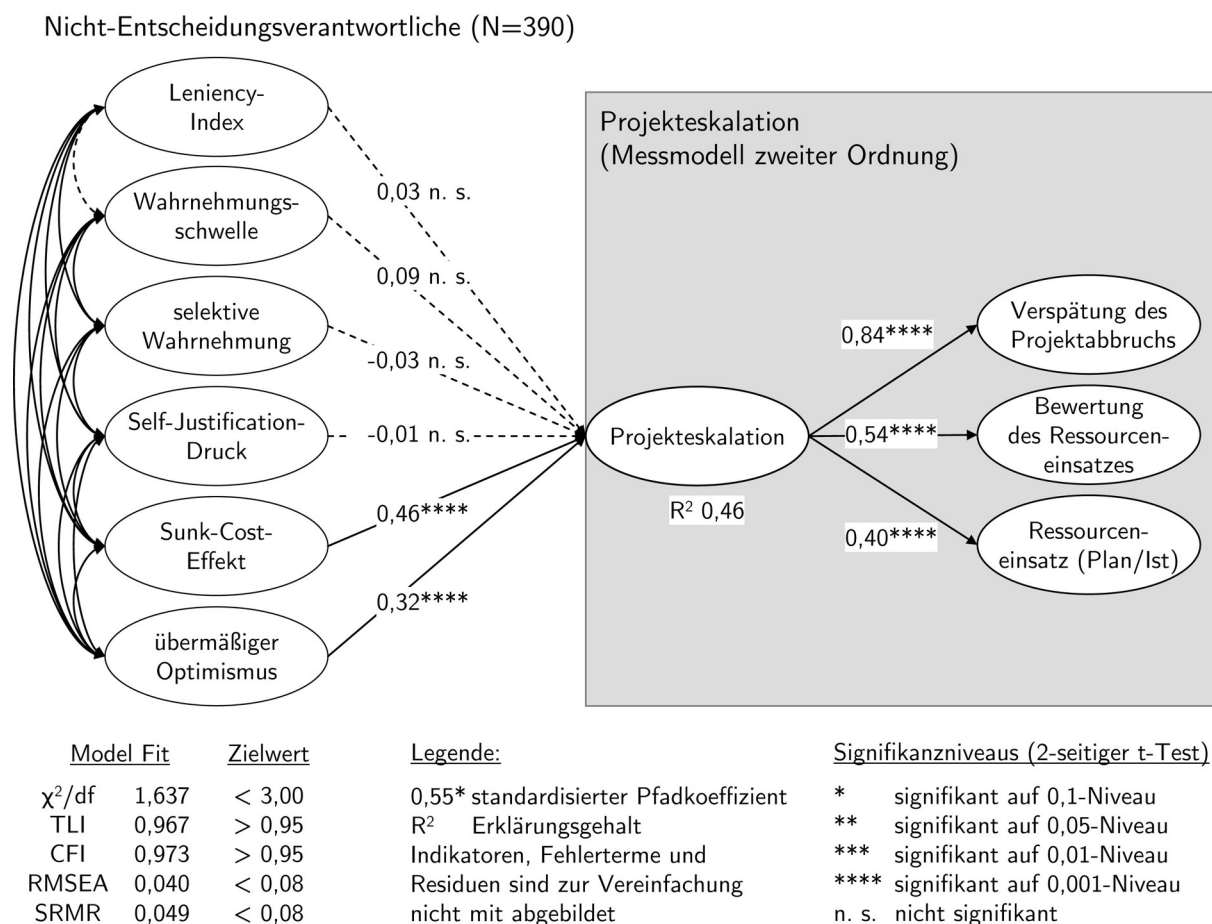
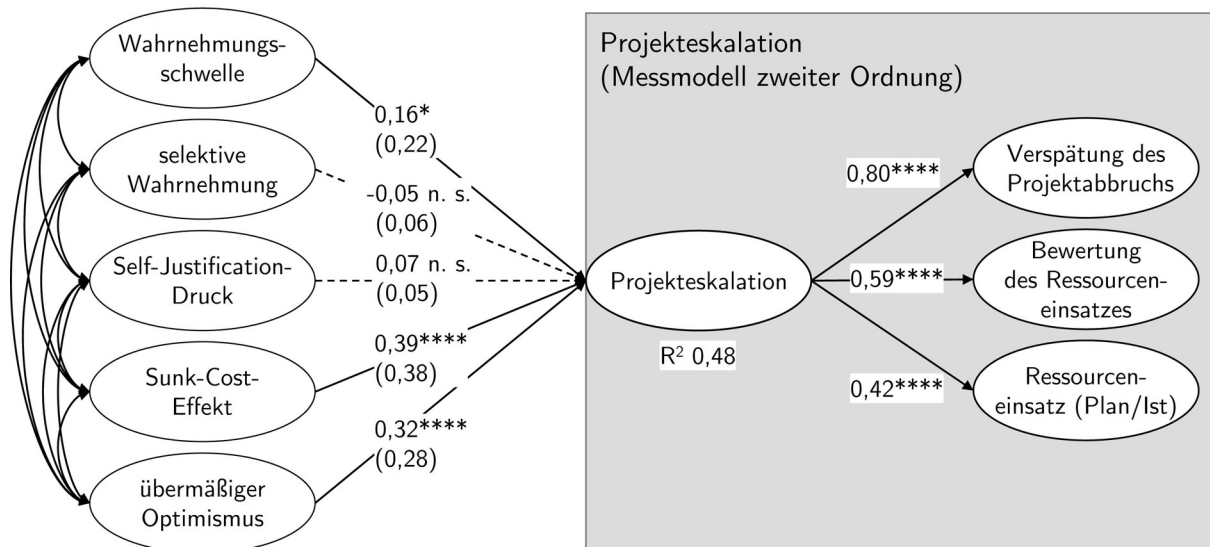


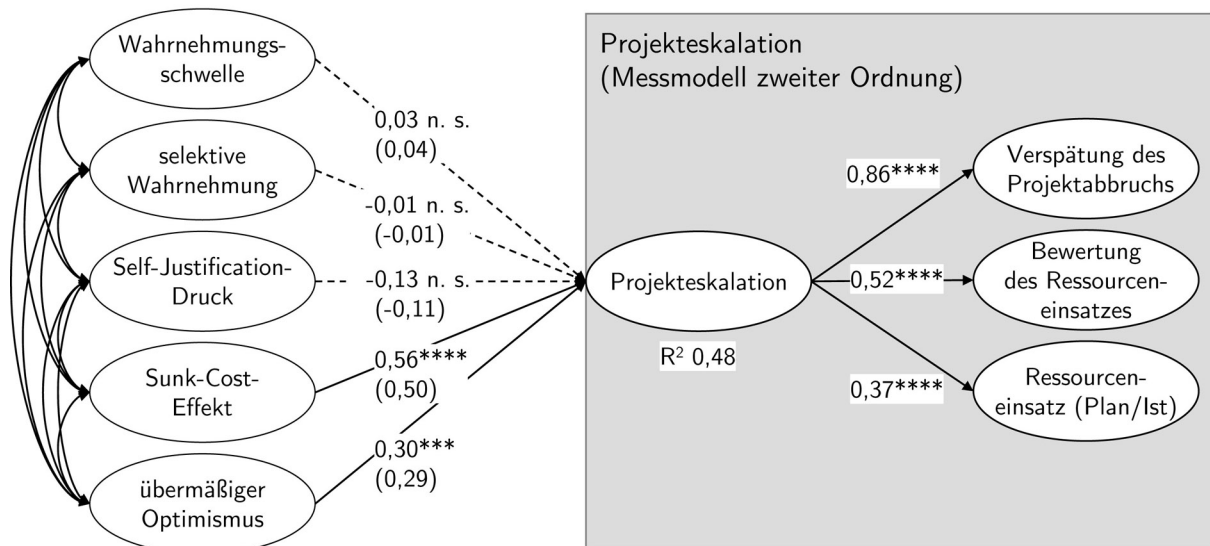
Abbildung 4.19: Eskalation (Modell 2) mit Leniency-Index

³³⁰ Vgl. Tabelle 4.19 auf Seite 140.

Gruppe 1 (N = 205): Leniency gering ausgeprägt



Gruppe 2 (N = 185): Leniency stark ausgeprägt



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df	1,449 < 3,00	0,55* standardisierter Pfadkoeffizient	* signifikant auf 0,1-Niveau
TLI	0,957 > 0,95	(0,55*) unstandardisierter Pfadkoeffizient	** signifikant auf 0,05-Niveau
CFI	0,963 > 0,95	R ² Erklärungsgehalt	*** signifikant auf 0,01-Niveau
RMSEA	0,034 < 0,08	Indikatoren, Fehlerterme und Residuen sind zur Vereinfachung nicht mit abgebildet	**** signifikant auf 0,001-Niveau
SRMR	0,058 < 0,08		n. s. nicht signifikant

Abbildung 4.20: Leniency als moderierender Faktor für Modell 2

Das Modell besitzt 227 Freiheitsgrade und einen χ^2 -Wert von 371,49. Alle Gütemaße erfüllen die geforderten Grenzwerte. Der Leniency-Index zeigt *keine* signifikante direkte Wirkung auf die Projekteskalation. Abgesehen von der Wahrnehmungsschwelle, weisen alle kognitiven Fak-

toren signifikante negative Korrelationen mit dem Leniency-Index auf.³³¹ Je größer also der Leniency-Effekt, desto weniger stark werden die kognitiven Faktoren bei dem Entscheidungsverantwortlichen beurteilt. Dies spricht für das Vorliegen einer Verzerrung der Ergebnisse durch den Leniency-Bias. Diese Verzerrung verändert jedoch *nicht* die Interpretation der Ergebnisse. Dies zeigt ein Vergleich zwischen Abbildung 4.18 auf Seite 190 und 4.19 auf Seite 191. Weder die standardisierten Pfadkoeffizienten noch der Erklärungsgehalt (R^2) weisen Unterschiede auf.

Um diese Schlussfolgerung zu prüfen, wurde zusätzlich ein Gruppenvergleich durchgeführt. Der Leniency-Index wurde dabei als moderierende Variable eingesetzt.³³² Die Ergebnisse der moderierenden Analyse sind in Abbildung 4.20 dargestellt. Das Modell besitzt 435 Freiheitsgrade und einen χ^2 -Wert von 630,107. Alle Gütemaße erfüllen die geforderten Grenzwerte. Die geschätzten Modelle sind für beide Gruppen ähnlich. Auffällig ist die Veränderung des Pfadkoeffizienten der Wahrnehmungsschwelle. Während er für die Gruppe 1 (mit geringem Leniency-Index) auf 0,1-Niveau signifikant ist, trifft dies in der Gruppe 2 nicht zu. Ein χ^2 -Differenztest zeigt jedoch, dass weder der Unterschied bei diesem Pfadkoeffizienten noch bei den anderen signifikant ist.

Die Resultate der expliziten Berücksichtigung des Leniency-Index stellen somit die in Abschnitt 4.4.1 ermittelten Ergebnisse nicht in Frage.

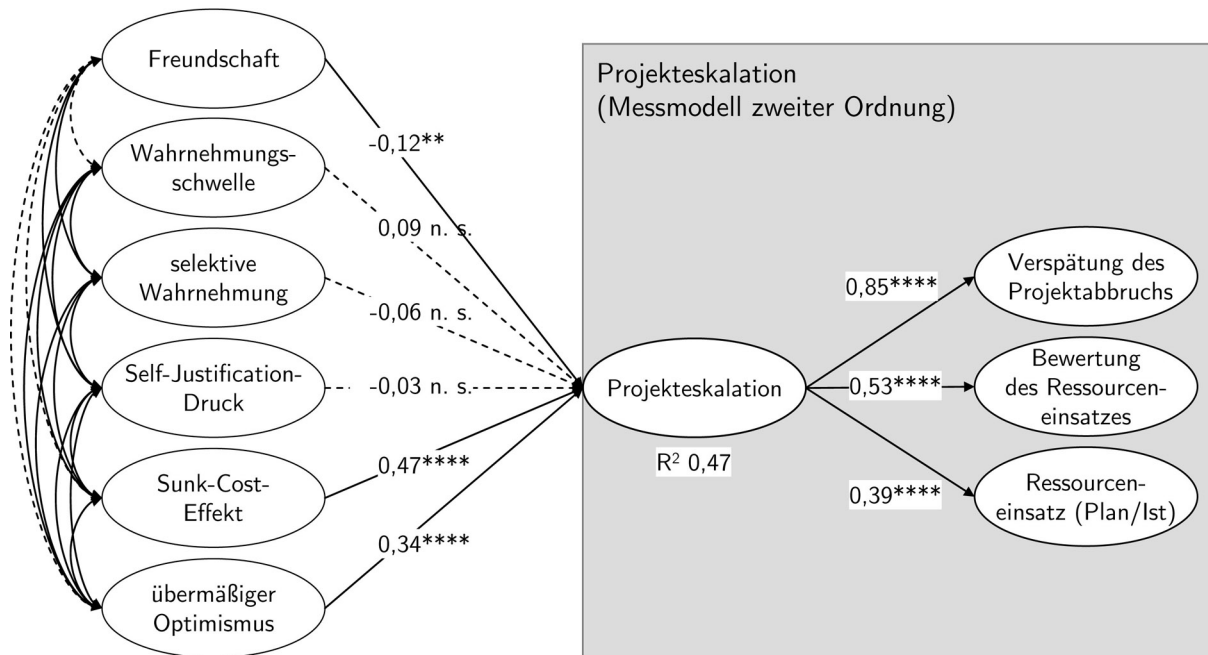
Freundschafts-Konstrukt

Beim Design des Fragebogens wurde noch eine weitere Möglichkeit implementiert, den Leniency-Bias zu prüfen. Diese Prüfung basiert auf der Annahme, dass Personen andere umso stärker in sozial erwünschter Weise beschreiben, je enger der Grad der freundschaftlichen Verbundenheit ist. Um zu prüfen, ob die Effekte der kognitiven Faktoren auf die Projekteskalation durch die Stärke der freundschaftlichen Beziehungen verzerrt werden, wurde deshalb ein Konstrukt zur Messung der Freundschaft verwendet. Der Wortlaut der Indikatoren und die Gütemaße wurden in Abschnitt 4.3.4.5 auf Seite 143 vorgestellt. Um den Effekt der Freundschaft zu prüfen, wird das Strukturmodell 2 (vgl. Abbildung 4.18 auf Seite 190) um das Freundschaftskonstrukt erweitert (vgl. Abbildung 4.21 auf der nächsten Seite). Da der Einfluss der Freundschaft nur für Nicht-Entscheidungsverantwortliche relevant ist, werden nur diese Respondenten in der Auswertung berücksichtigt.

³³¹ Die Korrelation des Leniency-Index mit der Wahrnehmungsschwelle beträgt -0,54 (t-Wert = -1,507; $p = 0,13$), mit der selektiven Wahrnehmung -0,44 (t-Wert = -7,388; $p < 0,0005$), mit dem Self-Justification-Druck -0,27 (t-Wert = -4,876; $p < 0,0005$), mit dem Sunk-Cost-Effekt -0,17 (t-Wert = -3,065; $p < 0,002$) und mit dem übermäßigen Optimismus -0,13 (t-Wert = -2,612; $p < 0,009$).

³³² Für eine Erläuterung der Vorgehensweise für moderierende Analysen vgl. Abschnitt 4.4.2 auf Seite 162.

Nicht-Entscheidungsverantwortliche (N=390)



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df 1,619	< 3,00	0,55* standardisierter Pfadkoeffizient	* signifikant auf 0,1-Niveau
TLI 0,966	> 0,95	R² Erklärungsgehalt	** signifikant auf 0,05-Niveau
CFI 0,971	> 0,95	Indikatoren, Fehlerterme und	*** signifikant auf 0,01-Niveau
RMSEA 0,040	< 0,08	Residuen sind zur Vereinfachung	**** signifikant auf 0,001-Niveau
SRMR 0,049	< 0,08	nicht mit abgebildet	n. s. nicht signifikant

Abbildung 4.21: Eskalation (Modell 2) mit Freundschafts-Konstrukt

Das Modell besitzt 273 Freiheitsgrade und einen χ^2 -Wert von 441,986. Alle Gütemaße erfüllen die geforderten Grenzwerte. Wie aus Abbildung 4.21 ersichtlich, weist die Freundschaft eine signifikante negative Wirkung auf die Projekteskalation auf. Die Beurteilung der Projekteskalation wird also in der erwarteten Weise durch den Grad der Freundschaft verzerrt: Je stärker die Freundschaft, desto geringer die Beurteilung der Projekteskalation.

Das Konstrukt Freundschaft zeigt signifikant negative Korrelationen zu selektiver Wahrnehmung und Self-Justification-Druck. Zu den verbleibenden drei kognitiven Faktoren sind die Korrelationen nicht signifikant.³³³

Vergleicht man das Modell ohne Freundschafts-Konstrukt mit dem Modell mit Freundschafts-Konstrukt, so lassen sich folgende Aussagen treffen: Für beide Modelle zeigen die selben ko-

³³³ Die Korrelation des Konstrukts Freundschaft mit der Wahrnehmungsschwelle beträgt -0,08 (t-Wert = -1,294; p = 0,20), mit der selektiven Wahrnehmung -0,15 (t-Wert = -2,603; p = 0,009), mit dem Self-Justification-Druck -0,10 (t-Wert = -1,737; p = 0,08), mit dem Sunk-Cost-Effekt -0,02 (t-Wert = -0,308; p = 0,76) und mit dem übermäßigen Optimismus 0,04 (t-Wert = 0,809; p = 0,42).

gnitiven Faktoren (Sunk-Cost-Effekt und übermäßiger Optimismus) eine signifikante Wirkung auf die Projekteskalation. Die Effekte sind also auch dann stabil, wenn die Freundschaft mit modelliert wird. Das R^2 erhöht sich durch das Freundschaftskonstrukt von 0,46 auf 0,47. Wird das Gesamtmodell um die Freundschaft ergänzt, verbessert sich der Erklärungsgehalt für die Projekteskalation also nur um 1 Prozent.

Nachdem geklärt wurde, dass die abhängige Variable Projekteskalation nur geringfügig durch den Grad der Freundschaft beeinflusst wird, stellt sich die Frage, ob die Wirkungen der kognitiven Faktoren einem Einfluss unterliegen. Um dies zu prüfen, wurde eine moderierende Analyse durchgeführt. Dazu wurden die Ausprägungen der Indikatoren des Freundschaftskonstrukts zu einem Index summiert. Anhand dieses Index wurden die Datensätze der Nicht-Entscheidungsverantwortlichen in zwei Gruppen eingeteilt. Gruppe 1 ($N = 205$) entspricht dabei einer geringen Ausprägung der Freundschaft, Gruppe 2 ($N = 185$) einer hohen. Die Übereinstimmung mit der moderierenden Analyse des Leniency-Index in Abschnitt 4.6.2.3 hinsichtlich der Anzahl der Datensätze je Gruppe ist zufallsbedingt. Abbildung 4.22 stellt die geschätzten Strukturgleichungsmodelle für die beiden Gruppen dar.

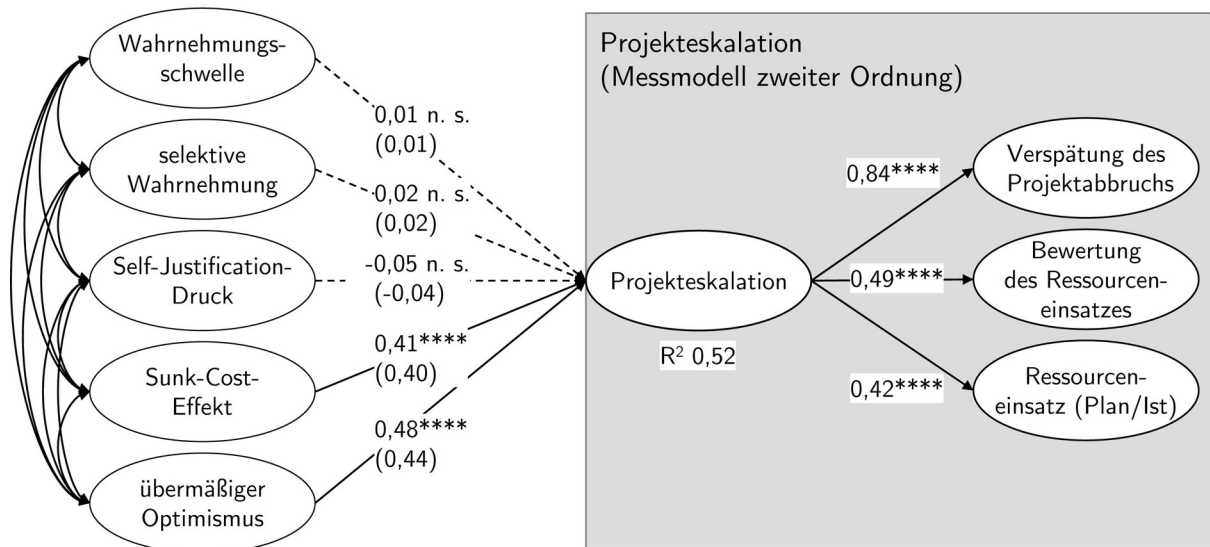
Das Modell besitzt 435 Freiheitsgrade und einen χ^2 -Wert von 590,139. Alle Gütemaße erfüllen die geforderten Grenzwerte. Die geschätzten Modelle sind für beide Gruppen ähnlich. Ein Blick auf die Pfadkoeffizienten zeigt allerdings zwei Unterschiede. (1) Bei geringer Freundschaft wirkt die Wahrnehmungsschwelle nicht signifikant auf die Projekteskalation, während bei stark ausgeprägter Freundschaft ein signifikanter Effekt auftritt. (2) Der standardisierte Pfadkoeffizient des übermäßigen Optimismus beträgt bei geringer Freundschaft 0,48 und ist auf 0,001-Niveau signifikant. Für die Gruppe mit starker Freundschaft beträgt der Pfadkoeffizient nur 0,15 und ist lediglich auf 0,1-Niveau signifikant.

Ein χ^2 -Differenztest zeigt,³³⁴ dass der unter (2) beschriebene Unterschied hinsichtlich des übermäßigen Optimismus signifikant ist ($\chi^2_{(1)} = 7,023$; $p = 0,008$). Die Unterschiede für alle anderen Pfadkoeffizienten, einschließlich der Wahrnehmungsschwelle, erweisen sich als nicht signifikant.

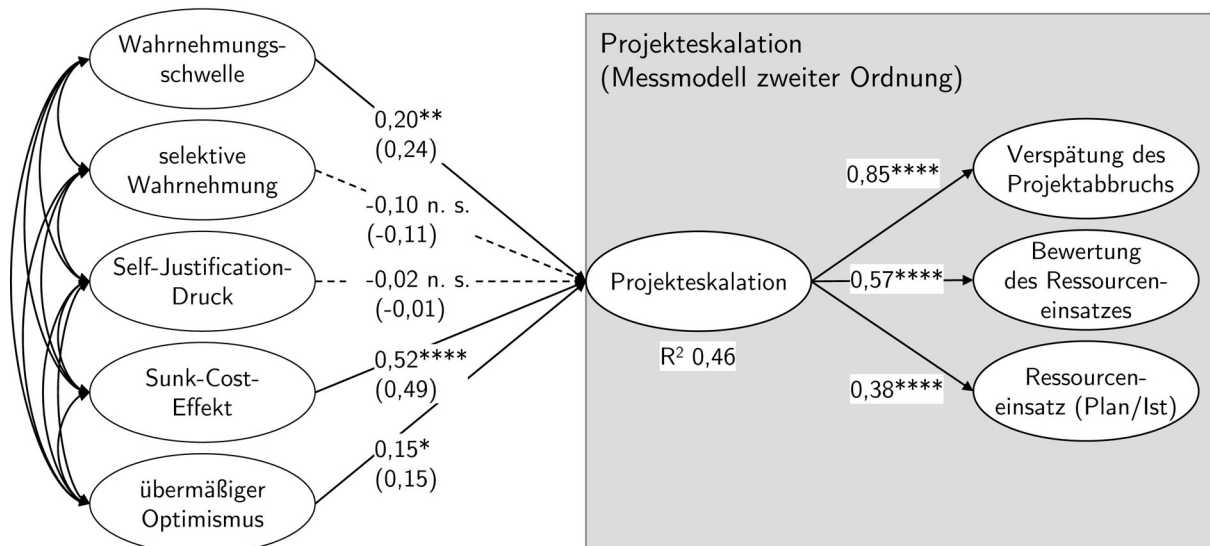
Die Berücksichtigung der Freundschaft führt somit nicht zu grundsätzlich anderen Interpretationen der Gesamtzusammenhänge als die Analysen in Abschnitt 4.4.1 auf Seite 148.

³³⁴ Für eine Erläuterung des χ^2 -Differenztests für die Untersuchung von moderierenden Effekten vgl. Abschnitt 4.4.2 auf Seite 162.

Gruppe 1 (N = 205): Freundschaft gering ausgeprägt



Gruppe 2 (N = 185): Freundschaft stark ausgeprägt



Model Fit	Zielwert	Legende:	Signifikanzniveaus (2-seitiger t-Test)
χ^2/df 1,417	< 3,00	0,55* standardisierter Pfadkoeffizient	* signifikant auf 0,1-Niveau
TLI 0,961	> 0,95	(0,55*) unstandardisierter Pfadkoeffizient	** signifikant auf 0,05-Niveau
CFI 0,966	> 0,95	R ² Erklärungsgehalt	*** signifikant auf 0,01-Niveau
RMSEA 0,033	< 0,08	Indikatoren, Fehlerterme und Residuen sind zur Vereinfachung nicht mit abgebildet	**** signifikant auf 0,001-Niveau
SRMR 0,056	< 0,08		n. s. nicht signifikant

Abbildung 4.22: Freundschaft als moderierender Faktor für Modell 2

Die Ergebnisse zeigen jedoch auch, dass der Leniency-Bias durchaus das Potential besitzt, die Ergebnisse signifikant zu verzerren. Würde beispielsweise das Gesamtsample nur aus Individuen bestehen, die eng mit dem Entscheidungsverantwortlichen befreundet sind, würde der Ef-

fekt des übermäßigen Optimismus weniger bedeutsam erscheinen (oder unter Umständen sogar die Signifikanzschwelle verfehlen), während er in einem gemischten Sample einen starken und deutlich signifikanten Erklärungsbeitrag für die Projekteskalation leistet.

4.6.2.4 Prüfung des Reihenfolge-Bias

Eine mögliche systematische Verzerrung der Ergebnisse besteht im Reihenfolge-Effekt.³³⁵ Insbesondere bei der Verwendung sehr vieler Items besteht die Gefahr, dass die Befragten nicht mehr sorgfältig zwischen den Fragen differenzieren, sondern ähnliche Urteile abgeben.³³⁶ Dies kann beispielsweise durch Ermüdung bzw. nachlassende Konzentration der Fall sein. Ein anderer möglicher Bias entsteht dadurch, dass die Befragten mentale Modelle für den Fragebogen entwickeln und versuchen, innerhalb ihres Modells die Fragen konsistent zu beantworten. Auch hierbei spielt die Reihenfolge der Fragen eine Rolle, da immer die bereits gelesenen Fragen die folgenden beeinflussen. Innerhalb jedes Frageblocks wurden die Items in einer variablen Zufallsfolge präsentiert.³³⁷

In Bezug auf die Reihenfolge der Fragebogenseiten wurden zwei verschiedene Anordnungen an jeweils die Hälfte der befragten Controller übersendet. Für die beiden Samples ($N_1 = 164$ und $N_2 = 141$) wurden t-Tests der Mittelwerte für alle Variablen im Fragebogen durchgeführt. Ein Reihenfolgeeffekt gilt als ausgeschlossen, wenn die arithmetischen Mittelwerte sich nicht signifikant unterscheiden.³³⁸ 15 der 76 metrisch skalierten Variablen weisen auf einem Signifikanzniveau von 0,05 signifikante Mittelwertunterschiede auf.³³⁹ Da die Anzahl der signifikanten Unterschiede fast viermal so hoch liegt, wie es beim Signifikanzniveau von 0,05 zu erwarten wäre, falls sich die Gruppen nicht unterschieden, scheint die Reihenfolge einen Einfluss auf die Ergebnisse zu haben. Die Effektstärke für t-Tests mit unabhängigen Stichproben lässt sich gemäß Formel 4.4 bestimmen.³⁴⁰

$$\eta^2 = \frac{t^2}{t^2 + (N_1 + N_2 - 2)} \quad (4.4)$$

³³⁵ Vgl. hierzu die Ausführungen in Abschnitt 3.3.4.3 auf Seite 73.

³³⁶ Vgl. Diller 2006, S. 616.

³³⁷ Vgl. Abschnitt 3.3.4.3 auf Seite 74.

³³⁸ Vgl. für diese Vorgehensweise beispielsweise Lueg/Finney 2007, S. 32.

³³⁹ Auf ein Verfahren zur Korrektur der Alphafehler-Kumulierung (auch als α -Fehler-Inflation bezeichnet) wurde an dieser Stelle verzichtet. Ein solches Verfahren, wie beispielsweise die Bonferroni-Korrektur, würde zu deutlich weniger signifikanten Unterschieden führen; vgl. Abdi 2007, S. 103 ff., und Pallant 2005, S. 247.

³⁴⁰ Vgl. Pallant 2005, S. 208.

Die Ergebnisse für η^2 können zwischen null und eins liegen und repräsentieren den Anteil der Varianz der abhängigen Variablen, die durch die unabhängige Variable (Gruppe) verursacht wird. Die Effektstärken liegen im vorliegenden Fall für alle betrachteten Variablen unter 0,04 und sind damit überwiegend als klein einzustufen.³⁴¹ Daraus kann gefolgert werden, dass die Reihenfolge zwar die Ergebnisse beeinflusst, jedoch in einem so geringen Ausmaß, dass bei den nachfolgenden Analysen nicht mit relevanten Verzerrungen zu rechnen ist.

4.6.3 Prüfung weiterer Biases

Zur Identifikation weiterer möglicher Biases wurde eine Reihe von Gruppenvergleichen durchgeführt, die im Folgenden kurz zusammengefasst werden.

Wie in Abschnitt 4.1.1.2 auf Seite 100 erläutert, wurden besonders auffällige Random-Responder aus dem Datensatz entfernt. Um eine mögliche Wirkung des Random Responding in den verbleibenden Datensätzen zu prüfen, wurde ein Gruppenvergleich hinsichtlich der Ausfülldauer durchgeführt. Darüber hinaus wurden für potentiell flüchtiges Ausfüllen die Antworten auf die Aussage „Für das untersuchte Thema finde ich die Länge des Fragebogens [„viel zu kurz“... „viel zu lang“].“ für einen Gruppenvergleich herangezogen. Eine Wirkung der Teilnahmemotivation wurde anhand der Antworten auf die Aussagen (1) „Ich hätte auch an der Umfrage teilgenommen, wenn es als Dankeschön nur die Auswertung der Ergebnisse gegeben hätte (und kein Buch, iPod-Verlosung etc.).“ und (2) „Die Themenstellung des Fragebogens war interessant.“ kontrolliert. Die Analysen ergaben keine auffälligen Gruppenunterschiede.

Auch die Erfahrung des Respondenten kann eine Quelle systematischer Abweichungen darstellen. Deshalb wurde das Gesamtsample anhand folgender Fragen für Gruppenvergleiche aufgeteilt: (1) „Wie viele Jahre Berufserfahrung hatten Sie zum Zeitpunkt des Projektabbruchs?“ (2) „Wie viele Jahre waren Sie in Ihrem Unternehmen zum Zeitpunkt des Projektabbruchs schon tätig?“ Ebenfalls untersucht wurde die Möglichkeit, dass eine Markierung der Benachrichtigungs-E-Mail durch den E-Mail-Provider als „Spam“ zu einem flüchtigeren und damit möglicherweise abweichenden Ausfüllen führt. Dazu wurden die Antworten auf die Fragen (1) „Die Einladungs-E-Mail zu dieser Umfrage war bei mir in der Betreffzeile als ‚Spam‘ markiert.“ und (2) „Die Einladungs-E-Mail zu dieser Umfrage habe ich im Spam-Ordner entdeckt.“ verwendet. In der Erhebung wurde auch abgefragt, ob ein einzelner Entschei-

³⁴¹ Cohen 1988 klassifiziert 0,01 als kleinen, 0,06 als mittleren und 0,14 als großen Effekt; zitiert nach Pallant 2005, S. 219.

dungsträger für die Abbruchentscheidung verantwortlich war oder ob mehrere Verantwortliche zuständig waren. Für die beiden Ausprägungen wurde ein Gruppenvergleich durchgeführt. Um den Informant-Bias zu prüfen,³⁴² wurden leitende Manager mit Controllern verglichen. Für alle genannten Kontrollgrößen traten keinen nennenswerten Effekte auf, so dass eine verzerrende Wirkung ausgeschlossen werden kann.

Einen Einfluss auf das Antwortverhalten könnte außerdem die Zugehörigkeit zum Unternehmen haben. Beispielsweise wäre es denkbar, dass diejenigen Personen, die noch im gleichen Unternehmen arbeiten, die Situation positiver darstellen, und Personen, die das Unternehmen gewechselt haben oder als Berater von außen beurteilen, kritischer zurückblicken. Deshalb wurde die folgende Frage für Gruppenvergleiche herangezogen: „Arbeiten Sie momentan für das Unternehmen, bei dem das Projekt stattfand?“ Die Antwortmöglichkeiten lauteten: „Ja, ich arbeite gegenwärtig für das Unternehmen, bei dem das Projekt stattfand.“ oder „Nein (z. B. wegen Jobwechsel oder weil sie als externer Berater in dem Unternehmen waren/sind).“ Bei dieser Prüfung zeigten sich für den Großteil der Variablen keine signifikanten Unterschiede. Dies spricht dafür, dass keine systematische Verzerrung (zum Beispiel eine Beschönigung aller Aspekte durch die Personen im gleichen Unternehmen) vorliegt.

Deutliche Differenzen gab es jedoch hinsichtlich des Self-Justification-Drucks. Der Gruppenvergleich zeigt, dass alle vier Items zur Messung des Self-Justification-Drucks auf einem Signifikanzniveau von 0,001 signifikante Mittelwertunterschiede aufweisen. Gleichzeitig handelt es sich bei genau diesen vier Items um die vier stärksten Effekte unter allen Items des Fragebogens. Bei allen vier Items liegt der Mittelwert bei den Personen, die nicht mehr im damaligen Unternehmen arbeiten, um 0,6 oder mehr (auf der siebenstufigen Skala) höher. Personen, die nicht mehr im gleichen Unternehmen arbeiten, geben also an, einen höheren Druck empfunden zu haben. Diese Unterschiede sind plausibel dadurch zu erklären, dass ein hoher empfundener Druck zum Wechsel des Arbeitgebers führen kann oder andersherum, in Fällen, wo bei Scheitern des Projekts auch eine Entlassung droht, der Druck besonders hoch ist. Diese Mittelwertunterschiede geben deshalb keinen Anlass, verfälschte Antworten zu vermuten.

Das untersuchte Sample umfasst Projekte mit sehr unterschiedlicher Größe, beispielsweise hinsichtlich der Investitionssumme bis zum Abbruchzeitpunkt.³⁴³ Um zu prüfen, ob die Projekt-

³⁴² Vgl. Abschnitt 3.3.3 auf Seite 65.

³⁴³ Vgl. Abbildung 4.1 auf Seite 108.

größe einen Einfluss auf die Wirkung der kognitiven Faktoren und der Gegenmaßnahmen hat,³⁴⁴ wurden moderierende Analysen durchgeführt.³⁴⁵ Für die kognitiven Faktoren ergaben die χ^2 -Differenztests keine signifikanten Unterschiede.³⁴⁶ Auch die Wirkung der Summe aller Maßnahmen³⁴⁷ auf die kognitiven Faktoren ist nicht von der Projektgröße abhängig.

Der einzige signifikante Unterschied liegt hinsichtlich der Wirkung der Summe aller Maßnahmen auf die Projekteskalation vor. Bei kleinen Projekten wirkt die Summe aller Maßnahmen nicht auf die Projekteskalation, während sie bei großen Projekten eine signifikante, negative (d. h. eskalationsreduzierende) Wirkung aufweist.³⁴⁸ Auf der Basis der vorliegenden Daten kann dieser Unterschied nicht zufriedenstellend erklärt werden, so dass weitere Untersuchungen zur Wirkungsweise der Maßnahmen notwendig sind, wie dies in Abschnitt 4.5.1.7 auf Seite 175 diskutiert und empfohlen wird.

4.6.4 Prüfung der Modellstabilität

Homburg/Klarmann empfehlen, insbesondere für komplexe Modelle mit mehr als zehn Konstrukten, Stabilitätstests durchzuführen.³⁴⁹ Zu diesem Zweck sollen mehrfach zehn Prozent der Fälle aus der Stichprobe zufällig entfernt werden. Für die vorliegende Untersuchung wurden deshalb drei Prüfsamples erstellt, bei denen jeweils zehn Prozent der Fälle zufällig gelöscht wurden. Mit einem χ^2 -Differenztest wurde geprüft, ob sich der Modell-Fit zwischen den Gruppen signifikant unterscheidet.³⁵⁰ Dies war nicht der Fall. Die Messung kann somit als stabil und der Stichprobenumfang als ausreichend groß bewertet werden.³⁵¹

³⁴⁴ Als Maß für die Projektgröße wurde die vor Projektbeginn geplante Gesamtinvestitionssumme herangezogen. Diese ergibt sich aus dem Quotient der bis zum Abbruch investierten Kosten und dem prozentualen Anteil der investierten Kosten im Vergleich zur Planung; vgl. hierzu den Fragebogen im Anhang in Abschnitt A.5 auf Seite 222.

³⁴⁵ Dazu wurde der Datensatz anhand der Projektgröße in zwei Gruppen aufgeteilt. Für eine Erläuterung zur Untersuchung von moderierenden Effekten vgl. Abschnitt 4.4.2 auf Seite 162.

³⁴⁶ Für eine Erläuterung des χ^2 -Differenztests vgl. Schermelleh-Engel/Moosbrugger/Müller 2003, S. 33 ff.

³⁴⁷ Vgl. Abbildung 4.12 auf Seite 177.

³⁴⁸ Dabei unterscheidet sich die Häufigkeit des Maßnahmeneinsatzes für fast alle Maßnahmen *nicht* signifikant zwischen kleineren und größeren Projekten.

³⁴⁹ Vgl. Homburg/Klarmann 2006, S. 728 und 737. Für ein Anwendungsbeispiel vgl. Homburg/Stock 2004. Für eine alternative Prozedur des Stabilitätstests vgl. beispielsweise Backhaus et alii 2006, S. 720 f.

³⁵⁰ Für die Forderung nach diesem Test zur Prüfung der Stabilität vgl. Backhaus et alii 2006, S. 720 f.

³⁵¹ Vgl. Homburg/Klarmann 2006, S. 733 ff.

5 Zusammenfassung und Fazit

5.1 Zentrale Ergebnisse

Das Phänomen der Eskalation des Commitments ist seit mehr als drei Jahrzehnten Gegenstand zahlreicher wissenschaftlicher Untersuchungen.¹ Vor allem die experimentelle Forschung konnte eine Reihe von kognitiven Faktoren identifizieren, die zu ökonomisch-irrationalem Festhalten an scheiternden Investitionen führen. Trotz der Forderungen in der Literatur, zum einen sophisticatedere Modelle empirisch zu prüfen und zum anderen die experimentellen Ergebnisse mit Daten aus der Unternehmenspraxis extern zu validieren, sind großzahlige Befragungen, die sich den kognitiven Faktoren der Eskalation widmen, bislang äußerst selten.² Die vorliegende Arbeit greift diese Forschungslücke auf und zielt darauf ab, die folgenden Forschungsfragen zu beantworten.

Forschungsfrage 1: Lässt sich die Wirkung der theoretisch und experimentell ermittelten kognitiven Faktoren auf die Projekteskalation empirisch im Unternehmenskontext bestätigen? Wenn ja, welche relative Einflussstärke weisen die verschiedenen kognitiven Faktoren auf die Projekteskalation auf?

Forschungsfrage 2: Lässt sich die Wirkung der verhaltenswissenschaftlich basierten Maßnahmen auf die kognitiven Faktoren empirisch im Unternehmenskontext bestätigen? Wenn ja, welche relative Einflussstärke weisen die verschiedenen Maßnahmen auf die kognitiven Faktoren auf?

¹ Vgl. Abbildung 2.1 auf Seite 13.

² Vgl. Abschnitt 2.1.2 auf Seite 12.

Um die Forschungsfragen zu beantworten, wurde auf das Eskalationsmodell von Zayer zurückgegriffen, das zum gegenwärtigen Zeitpunkt die umfassendste Synthese der bisherigen Forschung darstellt.³ Darauf aufbauend wurden zunächst die kognitiven Faktoren und die Projekteskalation operationalisiert. Dazu wurde eine umfangreiche Recherche in der Literatur sowie in psychologischen Testdatenbanken vorgenommen. Zusätzlich wurden Interviews und Vorstudien zur Konstruktentwicklung (Sorting-Pretest, Rating-Pretest) durchgeführt. Anschließend wurde eine branchenübergreifende Online-Befragung durchgeführt, an der sich 677 Controller und technische Leiter mit Projekterfahrung beteiligten.

Mit der Befragung von Experten aus der Praxis (anstelle der meist studentischen Samples bisheriger experimenteller Forschung) liegt erstmals ein großzählig-empirischer Nachweis der Wirkung der kognitiven Faktoren auf die Eskalation im realen Unternehmensumfeld vor. Von den fünf untersuchten Faktoren wirken die drei Faktoren Wahrnehmungsschwelle, Sunk-Cost-Effekt und übermäßiger Optimismus kontextübergreifend signifikant auf die Projekteskalation. Der Self-Justification-Druck übt einen signifikanten *direkten* Effekt auf die Projekteskalation aus, wenn der nicht hinterfragte Freiraum des Entscheiders hoch ist. Der Effekt der selektiven Wahrnehmung ist hingegen zu schwach ausgeprägt, um die Signifikanzgrenzen zu überschreiten.

Erstmals wird eine Abschätzung der relativen Wirkungsstärke der verschiedenen Faktoren möglich. Unabhängig von der Wirkung der Faktoren untereinander, für die verschiedene Modelle denkbar sind, übt stets der Sunk-Cost-Effekt den größten direkten Einfluss auf die Projekteskalation aus. An zweiter Stelle steht der übermäßige Optimismus. Den drittstärksten direkten Effekt bewirkt die Wahrnehmungsschwelle.

Darüber hinaus wurde auf der Basis theoretischer Überlegungen eine Wirkung des Self-Justification-Drucks auf die anderen kognitiven Faktoren für plausibel erachtet. Die aus der Theorie abgeleiteten Wirkungen weisen einen guten Fit mit den empirischen Daten auf. Möglich sind jedoch auch andere Wechselwirkungen zwischen den kognitiven Faktoren. In einem exploratorischen Verfahren wurde ein Modell bestimmt, das mit den Daten einen besonders hohen Fit aufweist. Die exploratorischen Analysen deuten darauf hin, dass auch die kognitiven Faktoren, die keine *direkte* Wirkung auf die Projekteskalation aufweisen, über andere Faktoren *indirekt* Einfluss auf die Projekteskalation nehmen können.

Zusammenfassend ist für die untersuchten kognitiven Faktoren festzustellen, dass ihre Wirkung auf die Projekteskalation auch im realen Unternehmenskontext bestätigt wird. Somit kann die

³ Vgl. Zayer 2007, S. 225.

erste Forschungsfrage auf der Basis der empirischen Befunde beantwortet und die bisherigen experimentellen Ergebnisse in ihrer externen Validität weitgehend untermauert werden.

In Bezug auf die zweite Forschungsfrage wurden folgende Erkenntnisse gewonnen. Während die Wirkung der kognitiven Faktoren auf die Eskalation kontextübergreifend (d. h. für das gesamte Sample) deutlich statistisch nachweisbar ist, trifft dies für die Wirkung der Gegenmaßnahmen weniger zu. Lediglich fünf der 21 untersuchten Maßnahmen erweisen sich für die Gesamtbetrachtung über alle Befragten als signifikant wirksam für die *Reduzierung* ihres jeweiligen kognitiven Faktors. Zwei weitere Maßnahmen wirken signifikant *verstärkend* auf den kognitiven Effekt, den sie reduzieren sollen. Dies bestätigt die in der Literatur bereits geäußerten Warnungen vor den möglichen nachteiligen Folgen eines unreflektierten oder unausgewogenen Maßnahmeneinsatzes.⁴

Von den verschiedenen kognitiven Faktoren scheint insbesondere der Sunk-Cost-Effekt nur sehr schwer durch Maßnahmen des Projektcontrollings reduzierbar zu sein. Während für alle anderen kognitiven Effekte mindestens eine der empfohlenen Maßnahmen signifikant reduzierend wirkte, war dies beim Sunk-Cost-Effekt nicht der Fall. Dies ist umso bedenkenswerter, als der Sunk-Cost-Effekt unter allen betrachteten kognitiven Faktoren den stärksten Einfluss auf die Projekteskalation aufweist.

5.2 Implikationen für die Praxis

Aus der vorliegenden Studie ergibt sich eine Reihe von Implikationen für die Unternehmenspraxis.

- (1) Zunächst einmal zeigen die empirischen Daten, dass die *experimentell* ermittelten kognitiven Eskalationsfaktoren auch in der Praxis nachweisbar und von hoher Relevanz sind.
- (2) Aus den relativen Wirkungsstärken lässt sich schlussfolgern, dass insbesondere den Bewertungsfehlern, d. h. dem Sunk-Cost-Effekt und dem übermäßigen Optimismus, stärkere Aufmerksamkeit zukommen sollte. Diese beiden Faktoren stellen die Haupttreiber für die Projekteskalation dar.
- (3) Aus den Analysen der verhaltensorientierten Controllingmaßnahmen lässt sich ableiten, dass insbesondere fünf Maßnahmen für die Mehrzahl der Projekte zur Reduzierung kognitiver Fehl-

⁴ Vgl. Zayer 2007, S. 215 ff.

leistungen beitragen können: das Festlegen von Meilensteinen, die Quantifizierung der Projektinformationen, der Schutz des Ansehens des Entscheidungsverantwortlichen im Falle eines Abbruchs, die Bekanntgabe der Ziele und das Ex-ante-Festlegen von Abbruchkriterien. Darüber hinaus kann die ermittelte durchschnittliche Häufigkeit des Maßnahmeneinsatzes in den am besten gesteuerten Projekten als Benchmark herangezogen werden.⁵ Weniger erfolgreich steuernde Unternehmen unterscheiden sich dabei in Bezug auf ihre Projektsteuerung von den erfolgreichereren insbesondere dadurch, dass sie seltener eindeutige Kriterien zur Beurteilung des Projekterfolgs verwenden, die Prozessqualität der Entscheidungen seltener beurteilen (sondern stärker nur auf die Ergebnisse einer Entscheidung schauen statt auf die Güte der Entscheidung zum damaligen Zeitpunkt), das Ansehen des Projektverantwortlichen weniger schützen und den Projektverlauf seltener in konkreten Zahlen abbilden.

(4) Aus der Tatsache, dass nur etwa ein Viertel der Maßnahmenvorschläge aus der Literatur eine kontextübergreifende Wirkung in der postulierten Art und Weise zeigt, ergibt sich, dass die Literaturempfehlungen keinesfalls unkritisch umgesetzt werden sollten. Bevor die Empfehlungen entsprechender Studien befolgt werden, muss stets hinterfragt werden, ob ihre Ergebnisse für den spezifischen Kontext angemessen sind.

(5) Aus den Ergebnissen folgt auch, dass ein pauschaler Einsatz aller denkbaren Maßnahmen nur eine Second-best-Lösung darstellt. Zwar wirkt der Einsatz *aller* 21 Maßnahmen in Summe signifikant eskalationssenkend, was in dieser Studie erstmals nachgewiesen werden konnte. Andererseits führt der Einsatz eines kontextspezifischen Bündels *ausgewählter* Maßnahmen zu noch besseren Resultaten und senkt gleichzeitig den Aufwand und die Komplexität des Controllings.

(6) Die für die vorliegende Studie entwickelten Messinstrumente können zukünftig in der Praxis verwendet werden, um eskalierendes Commitment in der Projektsteuerung zu erkennen. Insbesondere können sie bei der Beurteilung helfen, welche kognitiven Faktoren besonders stark auftreten und dadurch die in Punkt (5) geforderte Anwendung ausgewählter kontextspezifischer Maßnahmen unterstützen. Dabei ist es von Vorteil, dass der Fragebogen sowohl für die Entscheidungsverantwortlichen als auch für Nicht-Entscheidungsverantwortliche validiert wurde. Die Messinstrumente können somit sowohl für Selbstauskünfte als auch für die Beurteilung des Entscheidungsverantwortlichen durch andere Personen verwendet werden.⁶

⁵ Vgl. Abbildung 4.13 auf Seite 180.

⁶ Für einen Fragebogen, der allgemeine Projektfaktoren betrachtet, um zwischen eskalierenden und nicht eskalierenden Projekten zu unterscheiden, vgl. Keil/Rai/Mann 2000, S. 650 ff.

Die Studie ermöglicht darüber hinaus eine grobe Abschätzung, in welcher Größenordnung sich gut gesteuerte von schlecht gesteuerten Projekten im Hinblick auf die bis zum Projektabbruch investierten Kosten unterscheiden. Bei den zehn Prozent der besser gesteuerten Projekte wurden bis zum Abbruch im Durchschnitt 51 Prozent der ursprünglich geplanten gesamten Projektkosten investiert. Die Projekte, in die nach Einschätzung der Respondenten *zu viel investiert* wurde, wurden im Durchschnitt erst abgebrochen, nachdem bereits 99 Prozent der ursprünglich geplanten Gesamtkosten investiert worden waren.⁷ Bei einem durchschnittlich geplanten Projektvolumen von 68 Mio. Euro macht dies eine durchschnittliche Differenz von knapp 33 Mio. Euro aus.

5.3 Implikationen für die Wissenschaft und weiterer Forschungsbedarf

Die vorliegende Untersuchung lässt sich der Forschungsströmung zuordnen, die Limitationen und mögliche Erweiterungen der klassischen ökonomischen Theorie untersucht.⁸ Speziell für das Phänomen der Eskalation des Commitments bestätigen die Ergebnisse der vorliegenden Arbeit, dass das beobachtete Phänomen mit *Erweiterungen* der Annahmen der klassischen ökonomischen Theorie umfassender erklärt werden kann. Frühere Untersuchungen können dahingehend präzisiert werden, dass insbesondere eine Erweiterung der Evaluating-Annahme eine hohe Erklärungskraft für die Projekteskalation aufweist. Die ökonomische Theorie kann das Auftreten der Projekteskalation also wesentlich genauer beschreiben, wenn die Akteure unter der Annahme modelliert werden, dass sie *nicht* immer in der Lage sind, Zustände korrekt in Bezug auf ihren Nutzen zu bewerten.

Weniger stark, aber dennoch statistisch nachweisbar, tragen zwei andere Erweiterungen zu einem verbesserten Erklärungsgehalt der klassischen ökonomischen Theorie bei: die Ergänzung der Resourceful-Annahme durch die Wahrnehmungsschwelle und der Expecting-Annahme durch den übermäßigen Optimismus.

⁷ In Abbildung 4.14 auf Seite 181 werden die Top 10 Prozent der Projekte *allen* verbleibenden Projekten gegenübergestellt. Letztere umfassen sowohl Projekte, in die nach Meinung der Respondenten zu viel investiert wurde, als auch solche, in die zu wenig investiert wurde. In diese Projekte wurden bis zum Abbruch durchschnittlich 88 Prozent der geplanten Kosten investiert.

⁸ Vgl. Abschnitt 2.1.1 auf Seite 9.

Im Gegensatz dazu leistet die Erweiterung der Resourceful-Annahme durch die selektive Wahrnehmung keinen direkten Erklärungsbeitrag für die empirisch beobachteten Daten zur Projekteskalation. Auch die Erweiterung der Expecting-Annahme durch die Overconfidence stellt keine Verbesserung für das Verständnis der Eskalation dar.

Aus forschungsmethodischer Perspektive ist die Übertragbarkeit der Laborergebnisse auf die Unternehmenswirklichkeit von hohem Interesse. Diesbezüglich ergeben sich zwiespältige Implikationen aus den vorliegenden Daten.

Im Hinblick auf die *kognitiven Faktoren* können die Ergebnisse bisheriger Forschungen weitgehend repliziert und somit die früheren Laborexperimente in ihrer externen Validität bestätigt werden. Die Mehrzahl der kognitiven Faktoren erweist sich auch im realen Unternehmensumfeld als relevant. Grundlegend kann die experimentelle Forschung somit als geeignete Methode für die Erforschung der kognitiven Eskalationsfaktoren *und* als aussagekräftig für Projekteskalationen in der realen Unternehmenswelt betrachtet werden.

Einschränkend ist jedoch zu bemerken, dass eine Betrachtung *einzelner* kognitiver Effekte (die der gängigen Praxis in der experimentellen Forschung entspricht) nicht ausreicht, um die Relevanz außerhalb des Labors beurteilen zu können. So wurde beispielsweise die selektive Wahrnehmung in Laborexperimenten als wirksamer Eskalationsfaktor identifiziert. Die Daten aus realen Projekten hingegen deuten an, dass die selektive Wahrnehmung im Vergleich zu den anderen untersuchten Faktoren vernachlässigbar für die Erklärung der Projekteskalation ist – zumindest hinsichtlich ihrer *direkten* Wirkung.

In Bezug auf die *Maßnahmen* zur Reduzierung der kognitiven Faktoren scheinen die experimentellen Ergebnisse früherer Studien weniger generalisierbar zu sein, als in Bezug auf die kognitiven Faktoren. Die Mehrzahl der Maßnahmen zeigt für den Großteil der Projekte nicht die konzeptionell erwartete Wirkung. Zukünftig sollten deshalb die Ergebnisse von experimentellen Studien noch stärker als bisher durch Informationen aus der realen Unternehmenswelt kreuzvalidiert werden, wenn Empfehlungen für die Praxis angestrebt werden.

Ebenfalls aus methodischer Sicht interessant für die zukünftige wissenschaftliche Forschung sind die entwickelten Messmodelle. Erstmals liegen validierte Konstrukte mit multiplen Indikatoren zur Messung der Projekteskalation und der kognitiven Faktoren „Wahrnehmungsschwelle“, „selektive Wahrnehmung“, „Self-Justification-Druck“, „Sunk-Cost-Effekt“ und „Optimismus“ vor. Auf diesen Messinstrumenten können zukünftige Umfragen im Forschungsfeld der Eskalation des Commitments aufbauen.

Darüber hinaus wurde eine verkürzte Version von Paulhus' Impression-Management-Subscale entwickelt.⁹ Dazu wurde in einer Vorstudie geprüft, welche der ursprünglichen Items zur heutigen Zeit im deutschsprachigen Kulturraum angemessen sind, um das Phänomen des Impression-Managements zu messen. Die Resultate dieser Untersuchung können, auch über das Themenfeld der Projekteskalation hinaus, in der weiteren empirischen Forschung Verwendung finden.

Basierend auf den Ergebnissen der Studie sollen im Folgenden Anregungen für die Ausrichtung zukünftiger Forschungsprojekte gegeben werden.

Die Untersuchung ergibt, dass die kognitiven Faktoren in realen Entscheidungsprozessen *nicht* vollkommen unabhängig voneinander auftreten. Die Wirkungsmechanismen zwischen den Faktoren sollten genauer erforscht werden. Dabei ist insbesondere von Interesse, inwiefern kognitive Faktoren, die nicht direkt auf die Eskalation wirken, möglicherweise indirekt über andere Faktoren einen Einfluss auf die Projekteskalation nehmen. Die exploratorischen Analysen der vorliegenden Arbeit deuten an, dass dies beispielsweise bei der selektiven Wahrnehmung der Fall sein kann.

In Bezug auf den kognitiven Faktor Optimismus stellt sich heraus, dass *übermäßiger* Optimismus als separates Konstrukt operationalisiert werden muss. In zukünftigen Studien sollte genauer untersucht werden, wie die verschiedenen Ausprägungen des Optimismus voneinander abgegrenzt werden können und welche Konsequenzen dies für Studien hat, die sich mit Ursachen und Wirkungen des Optimismus beschäftigen.

Die vorliegende Arbeit zeigt, dass die Maßnahmen zur Reduzierung der kognitiven Eskalationsfaktoren in Summe eine signifikante Reduzierung der Projekteskalation bewirken. Gleichzeitig ist jedoch festzustellen, dass die Mehrzahl der untersuchten Maßnahmen keine Wirkung auf den *speziellen* kognitiven Faktor zeigt, für dessen Reduzierung sie in der Literatur empfohlen wird. Besonders betroffen ist der Sunk-Cost-Effekt, der den stärksten direkten Effekt auf die Eskalation ausübt. Für keine der vorgeschlagenen Maßnahmen kann eine reduzierende Wirkung auf diesen Faktor nachgewiesen werden. Es besteht somit dringender Forschungsbedarf, kontextübergreifend wirksame Maßnahmen zur Reduzierung des Sunk-Cost-Effekts zu entwickeln.

Auch für die anderen kognitiven Faktoren stellt sich die Situation nur wenig zufriedenstellender dar. So erweist sich nur etwa jede vierte Maßnahme als kontextübergreifend wirksam. Der Erforschung kontextspezifisch geeigneter Maßnahmenbündel kommt damit in der Zukunft eine

⁹ Vgl. Paulhus 1986.

große Bedeutung zu. Als Kontext ist hierbei nicht nur die Situation des Unternehmens aufzufassen,¹⁰ sondern auch die individuellen Merkmale der beteiligten Akteure.¹¹

Vielpersprechend für die zukünftige Eskalationsforschung ist es deshalb, den Einfluss der Persönlichkeitsmerkmale der am Entscheidungsprozess Beteiligten zu untersuchen.¹² Dazu gibt es bereits eine Reihe von Arbeiten,¹³ die in das vorliegende Modell integriert und empirisch geprüft werden könnten. Drei vielversprechende Ansätze sollen kurz genannt werden.

(1) Smith/Keil stellen Persönlichkeitsmerkmale zusammen, die einen Einfluss auf die Unwilligkeit haben, schlechte Neuigkeiten zu berichten.¹⁴ Solche Persönlichkeitsmerkmale sind zweifellos wichtig für die Eskalationstendenz. Für Unternehmen erscheint es dabei potentiell lohnenswert, die von Smith/Keil vorgeschlagenen Persönlichkeitsmerkmale bei der Auswahl von Projektverantwortlichen zu bedenken. Im Rahmen der weiteren Forschung wäre es deshalb interessant, die konzeptionellen Überlegungen von Smith/Keil empirisch zu prüfen.

(2) Wong untersucht die Verbindung zwischen Risikofreude, Risikowahrnehmung und Ergebniserwartung mit der Eskalation des Commitments.¹⁵ Dabei stellt sich heraus, dass Risikofreude und Ergebniserwartung einen positiven Zusammenhang zur Eskalation des Commitments aufweisen. Diese Merkmale könnten mit dem vorliegenden Modell kombiniert und auf die daraus resultierenden Wirkungen hin untersucht werden.

(3) Meglino/Korsgaard stellen die These auf, dass die Neigung, sich um andere zu sorgen bzw. sich für andere zuständig zu fühlen (other orientation), zu einer Reduzierung der Eskalation des Commitments führt.¹⁶ Sie untermauern diese These durch die Ergebnisse mehrerer Studien zum Thema Eskalation des Commitments,¹⁷ die konsistent mit dieser Aussage sind.¹⁸ Auch dieses Merkmal könnte mit dem vorliegenden Modell kombiniert werden, um das Bild der kausalen Beziehungen zu erweitern.

¹⁰ Vgl. für kontingenztheoretische Erwägungen auf organisationaler Ebene beispielsweise Burns/Stalker 1961, Lawrence/Lorsch 1967, Thompson 1967 und Galbraith 1973.

¹¹ Vgl. hierzu die Beiträge der kognitiven Psychologie, beispielsweise Newell/Simon 1972 und Tversky/Kahneman 1974.

¹² Für die Untersuchung verschiedener Ebenen (beispielsweise *individueller* Persönlichkeit und *organisationaler* Merkmale) in der Accounting-Forschung vgl. Luft/Shields 2003, S. 195 ff.

¹³ Vgl. beispielsweise Kisfalvi 2000, Meglino/Korsgaard 2004 und Wong 2005.

¹⁴ Vgl. Smith/Keil 2003, S. 80 ff.

¹⁵ Vgl. Wong 2005.

¹⁶ Vgl. Meglino/Korsgaard 2004, S. 950.

¹⁷ Vgl. Harrison et alii 1999, Moon 2001b und Sharp/Salter 1997.

¹⁸ Vgl. Meglino/Korsgaard 2004, S. 951.

Im Zusammenhang mit dem Einfluss von Persönlichkeitsmerkmalen wäre auch eine Untersuchung interessant, inwiefern sich Männer und Frauen im Eskalationsverhalten in realen Unternehmenssituationen unterscheiden. Experimentelle Studien gelangen beispielsweise zu der Schlussfolgerung, dass Frauen sich weniger von ihren vergangenen Erfolgen oder Misserfolgen beeinflussen lassen,¹⁹ dafür jedoch stärker auf Konflikte (wie durch die *Advocatus-Diaboli*-Maßnahme)²⁰ reagieren.²¹

Neuere Forschungsansätze betrachten die Projekteskalation unter dem Gesichtspunkt der Realloptionentheorie.²² Dies eröffnet eine neue Perspektive auf das Phänomen der Eskalation, die dazu beitragen kann, das Auftreten kognitiver Fehlleistungen kritisch zu hinterfragen. Es steht zu erwarten, dass einige Beobachtungen, die auf den ersten Blick als Eskalationsphänomene erscheinen, *nicht* als ökonomisch irrational zu interpretieren sind, wenn der Wert der Realoptionen mit berücksichtigt wird.

Neben den kognitiven Faktoren, die in dieser Arbeit untersucht wurden, können auch physiologische Faktoren (d. h. biochemische Prozesse) einen Einfluss auf Eskalation ausüben. Es ist beispielsweise davon auszugehen, dass Euphorie (bzw. die dabei ausgeschütteten Hormone wie Adrenalin und Endorphin) die *Overconfidence* steigern.²³ Eine einfache, aus dieser Erkenntnis abzuleitende Maßnahme ist, dass in euphorischen Situationen keine Entscheidungen über Projektfortführungen getroffen werden sollten.²⁴ Zukünftige Forschungen in dieser Richtung bieten das Potential, die Wirtschaftswissenschaften durch Erkenntnisse aus den Neurowissenschaften zu bereichern, nachdem dies mit der Psychologie in den letzten Jahrzehnten bereits in Ansätzen gelungen ist.

5.4 Limitationen der Untersuchung

Ein grundsätzlich limitierender Faktor besteht in der schwierigen Messung kognitiver Phänomene in der Unternehmenspraxis.²⁵ So betont Camerer beispielsweise die besonderen Schwie-

¹⁹ Vgl. Bateman 1986.

²⁰ Vgl. hierzu den Abschnitt 2.3.4.2 auf Seite 41.

²¹ Vgl. Schwenk 1988, S. 780.

²² Vgl. beispielsweise Fichman/Keil/Tiwana 2005.

²³ Vgl. Russo/Schoemaker 1992, S. 15.

²⁴ Vgl. Russo/Schoemaker 1992, S. 15.

²⁵ Für die vorliegende Untersuchung trifft die Aussage von Armstrong/Williams/Barrett 2004 zu: "The measure-

rigkeiten, die mit der Erhebung und Auswertung von Daten zum Auftreten von Rationalitätsdefiziten außerhalb des Labors verbunden sind, und fordert deshalb besonders strikte Prüfungen und die Anwendung von Serien von Studien: “First, establishing systematic mistakes using naturally occurring data is very difficult. Of course, this does not mean we should avoid such hard work and exploit the superior control of the lab; it just means that the standard of proof for mistakes outside the lab is high, and should be. It is also likely that important field anomalies will not be established by a single study, but by a series of studies which build on earlier results.”²⁶ Dies gilt auch für die vorliegende Untersuchung. In methodischer Hinsicht stellen die entwickelten Messmodelle latenter Konstrukte einen Schritt nach vorn dar. Nachfolgende Studien sind jedoch essentiell, um die Validität der Messungen zu prüfen und die gefundenen Aussagen weiter differenzieren zu können.

Eine Einschränkung bildet der Umfang des Modells. Wie in der Herleitung dargestellt,²⁷ basiert das Modell notwendigerweise auf einer Reihe vereinfachender Annahmen. Eine starke Einschränkung stellt beispielsweise der Fokus auf die *kognitiven* Faktoren *eines* Entscheidungsverantwortlichen dar. Dementsprechend bleiben alle nicht-kognitiven Einflussfaktoren unberücksichtigt. Ebenso konnte die Untersuchung weder den Einfluss von Gruppeneffekten noch das Zusammenspiel verschiedener Hierarchie-Level (einschließlich der damit verbundenen Informationsasymmetrien etc.) abbilden. Obwohl diese und weitere Elemente mit Sicherheit relevant für die Eskalation von Projekten sind, mussten sie aus Komplexitätsgründen ausgespart bleiben.

Des Weiteren beschränkt sich die Studie auf Erklärungsansätze, die schon zuvor erfolgreich in Experimenten nachgewiesen wurden. Außerdem existiert noch eine Reihe weiterer Erklärungsansätze, die möglicherweise in zukünftigen Forschungsprojekten empirisch belegt werden und dann auf ihre Kompatibilität mit dem getesteten Modell geprüft werden müssen. Dies gilt insbesondere für die Maßnahmen des verhaltensorientierten Controllings. In der Literatur werden noch zahlreiche andere Maßnahmen für die Reduzierung der Eskalation des Commitments vorgeschlagen. Die Auswahl fiel bewusst auf die von Zayer empfohlenen Maßnahmen, da diese einen direkten Bezug zu den verwendeten Theorien aufweisen. Dennoch ist es nicht auszuschließen, dass andere, nicht berücksichtigte Maßnahmen einen deutlich stärkeren eskalationsreduzierenden Effekt haben. Überdies werden einige der hier untersuchten Maßnahmen in

ment process in this study reflects a bold attempt to measure variables that are very difficult to measure with primary data.”; Armstrong/Williams/Barrett 2004, S. 369.

²⁶ Camerer 1999, S. 81.

²⁷ Vgl. Abschnitt 2.2 auf Seite 21.

der Literatur nicht für die Reduzierung der gleichen kognitiven Faktoren empfohlen, denen sie in der vorliegenden Studie zugeordnet wurden; stattdessen wird ihnen eine Wirkung auf andere kognitive Faktoren zugeschrieben.²⁸

Die Vorgehensweise, nach vergangenen Ereignissen zu fragen, birgt das Risiko des Hindsight-Bias – die Respondenten schätzen das Projekt möglicherweise im Nachhinein anders ein, als es sich während der Laufzeit darstellte.²⁹

Schließlich limitiert der Fokus auf gescheiterte Projekte die Interpretation. Für zukünftige Studien könnte es interessant sein, den Maßnahmeneinsatz und den Einfluss der kognitiven Faktoren auch in erfolgreich abgeschlossenen Projekten zu untersuchen, um Vergleiche zwischen erfolgreichen und gescheiterten Projekten durchführen zu können.

²⁸ Vgl. Absatz 4.5.1.7 auf Seite 175.

²⁹ Vgl. Abschnitt 3.3.2 auf Seite 63.

Anhang

A.1 Vorstudie zum Impression-Management: Einleitungstext

Auf dieser Seite geht es darum zu beurteilen, wie bestimmte persönliche Fragen in Deutschland empfunden werden. Aus Ihren Aussagen sollen Schlüsse gezogen werden, inwieweit die Einzelaussagen in der Hauptbefragung verwendet werden können.

Bei den folgenden Fragen geht es darum zu beurteilen, wie brisant die einzelnen Verhaltensweisen in Deutschland gesehen werden. D. h., wenn jemand bei anderen einen guten Eindruck machen will, wie wichtig wäre es, dass er bei folgenden Aussagen ggf. lügt, um nicht zugeben zu müssen, die unten genannten Dinge getan zu haben?

Bitte stellen Sie sich dazu Standardsituationen aus Ihrem Leben vor (z. B. ein Gespräch mit Chef, Kollegen etc.). Angenommen, das Gespräch käme auf die folgenden Themen. Wie wichtig wäre es, dass Sie bestimmte Verhaltensweisen leugnen, damit die anderen keinen schlechten Eindruck bekommen?

Beispiel: Im Gespräch mit Kollegen kommt das Thema auf zu schnelles Autofahren. Angenommen, Sie würden gelegentlich zu schnell fahren, wäre es dann wichtig, dies zu verheimlichen, damit die anderen keinen schlechten Eindruck gewinnen? Wenn es völlig unwichtig wäre, bei diesem Thema die Wahrheit zu verheimlichen, kreuzen Sie „völlig unwichtig“, wenn es extrem wichtig wäre, kreuzen Sie „extrem wichtig“ an.

Achtung, bitte NICHT antworten, wie sehr die Aussagen auf Sie zutreffen, sondern bewerten, wie wichtig es in Deutschland ist, bei den folgenden Themen eigene Fehler zu verheimlichen.

Tabelle A.1: Impression-Management-Subscale

Nr.	Originalindikatoren (englisch)	Deutsche Übersetzung
1.	Sometimes I tell lies if I have to.	Manchmal erzähle ich Lügen, wenn ich muss.
2.	I never cover up my mistakes.	Ich vertusche meine Fehler nie.
3.	There have been occasions when I have taken advantage of someone.	Es gab Gelegenheiten, bei denen ich jemanden ausgenutzt habe.
4.	I never swear.	Ich fluche nie.
5.	I sometimes try to get even rather than forgive and forget.	Manchmal versuche ich etwas heimzuzahlen, statt zu vergeben und zu vergessen.
6.	I always obey laws, even if I'm unlikely to get caught.	Ich befolge Gesetze immer, auch wenn es unwahrscheinlich ist, dass ich erwischt werde.
7.	I have said something bad about a friend behind his/her back.	Ich habe schon Schlechtes über einen Freund hinter seinem Rücken gesagt.
8.	When I hear people talking privately, I avoid listening.	Wenn ich höre, wie Leute sich privat unterhalten, versuche ich nicht hinzuhören.
9.	I have received too much change from a salesperson without telling him or her.	Ich habe schon zu viel Wechselgeld von einem Verkäufer erhalten, ohne es zu sagen.
10.	I always declare everything at customs.	Ich verzolle immer alles, was ich verzollen muss.
11.	When I was young, I sometimes stole things.	Als ich jung war, habe ich manchmal gestohlen.
12.	I have never dropped litter on the street.	Ich habe noch nie Abfall auf die Straße geworfen.
13.	I sometimes drive faster than the speed limit.	Ich fahre manchmal schneller als erlaubt.
14.	I never read sexy books or magazines.	Ich lese niemals erotische Bücher oder Magazine.
15.	I have done things that I don't tell other people about.	Ich habe schon Dinge getan, die ich niemandem erzähle.
16.	I never take things that don't belong to me.	Ich nehme nie Dinge, die mir nicht gehören.
17.	I have taken sick leave from work or school even though I wasn't really sick.	Ich habe mich auf der Arbeit oder in der Schule schon einmal krank gemeldet, obwohl ich nicht wirklich krank war.
18.	I have never damaged a library book or store merchandise without reporting it.	Ich habe noch nie ein Buch in der Bücherei oder einen Handelsartikel beschädigt, ohne es zu sagen.
19.	I have some pretty awful habits.	Ich habe einige ziemlich schreckliche Angewohnheiten.
20.	I don't gossip about other people's business.	Ich tratsche nicht über anderer Leute Angelegenheiten.

Im Pretest wurden nur die deutschen Formulierungen verwendet.

Die englischen Originalitems sind hier als Zusatzinformation angegeben, um die Übersetzung beurteilen zu können.

[Die Antwortmöglichkeiten bestanden in einer siebenstufigen Skala von „völlig unwichtig“ bis „extrem wichtig“.]

In diesem Block geht es darum zu beurteilen, wie beleidigend die Aussagen wirken. Aussagen, die beleidigend wirken, wollen wir in unserem Fragebogen vermeiden. Stellen Sie sich deshalb vor, Sie würden an der Hauptbefragung zur „Projektsteuerung“ teilnehmen. Wie sehr würden Sie sich beleidigt fühlen und vielleicht sogar den Fragebogen abbrechen, wenn Sie zu den folgenden Aussagen Stellung beziehen müssten?

Achtung, bitte NICHT antworten, wie sehr die Aussagen auf Sie zutreffen, sondern bewerten, wie beleidigend Sie es finden würden, wenn Sie in einem Fragebogen Stellung zu den Aussagen nehmen müssten.

[An dieser Stelle folgten die zwanzig Items der Impression-Management-Skala in deutscher Sprache. Tabelle A.1 auf der vorherigen Seite gibt die Indikatoren im Wortlaut wieder und enthält zusätzlich zum Vergleich die englischen Originalformulierungen. Die Antwortmöglichkeiten bestanden in einer siebenstufigen Skala von „gar nicht beleidigend“ bis „extrem beleidigend“. Außerdem folgte ein Freitextfeld mit der Erläuterung: „Ihre Kommentare zu diesen Aussagenblöcken“.]

A.2 Sorting-Pretest: Begrüßungstext und Konstruktdefinitionen

Herzlich willkommen zum Pre-Pretest für meine empirische Untersuchung. In meinem Forschungsprojekt geht es um den Einfluss kognitiver Begrenzungen auf die Abbruchentscheidungen bei scheiternden Projekten. Der Pretest umfasst zwei Seiten. Auf der ersten Seite geht es darum, zu beurteilen, inwieweit bestimmte Fragen aus dem amerikanischen Raum auf unseren Kulturkreis übertragbar sind. Auf der zweiten Seite wird untersucht, wie gut bestimmte Einzelfragen verschiedenen vordefinierten Oberklassen zuzuordnen sind. Das Ausfüllen des gesamten Fragebogens dauert ca. 15 Minuten. Für Rückfragen steht ich Dir gern zur Verfügung. Herzlichen Dank!!!

Bitte lies zunächst die Definitionen auf der Seite aufmerksam durch.¹ Ordne dann die einzelnen Fragen den Kategorien zu. Wenn keine der Kategorien passend erscheint, markiere bitte die letzte Spalte „Sonstiges“.

Wahrnehmungsdefizit: das Ausmaß, in dem relevante Projektinformationen der Aufmerksamkeit des Entscheiders entgingen oder erst zu spät bemerkt wurden; d. h., inwieweit sind tatsächlich Probleme zu spät oder gar nicht bemerkt worden.

Wahrnehmungsschwelle: das Ausmaß, in dem negative Informationen nicht auffällig bzw. entscheidungsrelevant genug schienen, um die Wahrnehmungsschwelle des Entscheiders zu überwinden; d. h. das Potential der Probleme, auf Grund der Wahrnehmungsschwelle nicht bemerkt zu werden.

Selektive Wahrnehmung: das Ausmaß, in dem der Entscheider dazu neigte, selektiv wahrzunehmen (d. h. in diesem Fall, negative Informationen auszublenden).

Individuelles Zielsystem: das Ausmaß, in dem der Entscheider das Projekt aus persönlichen Gründen (z. B., um sein Gesicht zu wahren) weiterführte, obwohl dies den ökonomischen Zielen des Unternehmens widersprach.

Self-Justification-Druck: das Ausmaß, in dem der Entscheider Druck empfand, mit der Projektfortführung sein Gesicht (vor sich selbst und vor anderen) wahren zu müssen.

¹ Die Definitionen wurden während der Entwicklungsphase der Studie kontinuierlich überarbeitet. Die hier aufgeführten Definitionen können deshalb leichte Abweichungen von Definitionen aufweisen, die an anderer Stelle in dieser Arbeit aufgeführt werden.

Rationale Bewertung: das Ausmaß, in dem das Projekt ökonomisch rational bewertet wurde.

Einfluss der Sunk Costs: das Ausmaß, in dem die Sunk Costs (d. h. die bereits investierten unwiederbringlichen Ressourcen) als relevant für die Projektbewertung empfunden wurden.

Optimismus: das Ausmaß, in dem der Entscheider an einen erfolgreichen Abschluss des Projekts geglaubt hat.

[An dieser Stelle folgten die Indikatoren (siehe Tabelle 3.7 auf Seite 85).]

Falls Du noch Kommentare, Anmerkungen etc. hast, bist Du herzlich eingeladen, diese hier zu notieren. Ich freue mich über jede Anregung! Interessant wäre auch eine kurze Erläuterung, wie Du bei der Aufgabe des Zuordnens vorgegangen bist. Also, ob Du z. B. irgendwelche gedanklichen Schemata zur Vereinfachung verwendet hast, um die Aufgabe zu lösen (also beispielsweise nach Schlüsselwörtern suchen, statt den gesamten Sinn der Aussage mit den Definitionen zu vergleichen o. Ä.) Ganz herzlichen Dank für Deine Unterstützung!

A.3 Rating-Pretest: Begrüßungstext

Vorstudie zur Befragung:

Erfolgreiche Projektsteuerung

Die deutschlandweit größte Umfrage zur Verbesserung von Projektabbruchentscheidungen

Von Ihrer Teilnahme hängt der Erfolg dieses Forschungsprojekts ab!

Nur durch die Auskünfte von Experten wie Ihnen können wir praxisbezogene Erkenntnisse gewinnen, die auch Ihnen zugute kommen.

Ihre Angaben werden anonym ausgewertet und streng vertraulich behandelt!

Die Ergebnisse werden ausschließlich in aggregierter Form veröffentlicht, so dass keine Rückschlüsse auf Sie persönlich oder Ihr Unternehmen möglich sind.

Das Ausfüllen des Fragebogens wird insgesamt etwa 15 Minuten dauern.

Als Dankeschön für Ihre Teilnahme erhalten Sie in jedem Fall:

- die Ergebnisse der in vier Wochen beginnenden Haupt-Studie in Form eines exklusiven Best-Practice-Berichtes mit einer Checkliste der erfolgreichsten Maßnahmen zur Verbesserung von Projektabbruchentscheidungen,

sowie zusätzlich wahlweise:

- kostenlos einen Band Ihrer Wahl der Bände 43, 44, 47, 50, 51 oder 52 aus der Reihe Advanced Controlling im Wert von 24,90 Euro

oder

- die Teilnahme an einer Verlosung mit der Chance von jeweils 1 zu 100, einen 8 GB iPod nano von Apple im Wert von 249 Euro zu gewinnen.

Für Rückfragen steht Ihnen Dipl.-Wi.-Ing. Matthias Mahlendorf gerne zur Verfügung:

Tel.: 0261/6509-482; Fax: 0261/6509-479

E-Mail: Projektcontrolling@whu.edu

Dieser Fragebogen ist ein Vortest für die eigentliche Hauptuntersuchung, die in einigen Wochen starten soll. Ziel dieses Vortests ist es, Informationen von Ihnen zu erhalten, die uns helfen, diejenigen Fragen auszuwählen, die für die Hauptbefragung am geeignetsten sind.

Im Folgenden bitten wir Sie zu beurteilen, wie gut einzelne Aussagen ein zuvor definiertes Konstrukt widerspiegeln. Ein Konstrukt ist dabei ein nicht direkt beobachtbarer Sachverhalt innerhalb einer wissenschaftlichen Theorie. Konstrukte ähneln somit Konzepten und sind gedanklicher bzw. theoretischer Natur. Das bedeutet nicht, dass der betreffende Sachverhalt nicht „existiert“, sondern nur, dass er aus anderen, leicht(er) beobachtbaren Sachverhalten (sog. Indikatoren) erschlossen wird. Beispiel: Das Konstrukt Intelligenz entzieht sich einer direkten Beobachtbarkeit, kann aber über einen Indikator wie den IQ näherungsweise gemessen werden. In unserem Fragebogen dienen also die Aussagen (Indikatoren) dazu, Rückschlüsse auf das definierte Konstrukt zu ziehen, das nicht direkt beobachtbar ist.

Bitte lesen Sie zunächst die Definition des Konstrukts durch. Bewerten Sie anschließend für jede Aussage im Block, ob sie das definierte Konstrukt „gar nicht repräsentiert“, „einigermaßen repräsentiert“ oder „klar repräsentiert“.

[An dieser Stelle folgte jeweils die Definition eines Konstrukts (vgl. Abschnitt A.2 auf Seite 216) sowie die Indikatoren zu dem entsprechenden Konstrukt (siehe Tabelle 3.7 auf Seite 85). Die Antwortmöglichkeiten lagen auf einer dreistufigen Skala von „gar nicht repräsentiert“ bis „klar repräsentiert“.]

A.4 Anschreiben

Sehr geehrte/r Frau/Herr XYZ,

die Entscheidung über Fortführung, Neustart oder Abbruch risikoreicher Projekte stellt eine der größten Herausforderungen für das technische Management dar. Zu diesem sicherlich auch für Sie relevanten Thema „Erfolgreiche Steuerung risikoreicher Projekte“ führt der Lehrstuhl für Controlling und Telekommunikation an der WHU (www.whu.edu/control) derzeit ein umfassendes Forschungsprojekt durch. Teil dieses Projekts ist eine Befragung von technischen Leitern aus verschiedenen Branchen. Wir möchten Sie bitten, durch Ausfüllen eines Online-Fragebogens zum Erfolg des Forschungsprojekts beizutragen.

Als Dankeschön für das Ausfüllen des Fragebogens (Dauer: ca. 20 Min.) erhalten Sie nach Auswertung der Ergebnisse folgende drei Leistungen:

- 1.) einen exklusiv erstellten Best-Practice-Bericht, basierend auf den Ergebnissen dieses Forschungsprojekts, aus dem Handlungsempfehlungen für Ihr Unternehmen ersichtlich werden,
- 2.) eine Checkliste mit aktuellen Maßnahmen zur Verbesserung der Projektsteuerung,
- 3.) ein Gratisexemplar einer Publikation unseres Lehrstuhls. Sie können wählen unter Titeln wie „Investitionscontrolling in deutschen Großunternehmen“, „Erfahrungen mit der Balanced Scorecard revisited“, „Strategisches Controlling“, „IFRS revisited: Quo vadis Unternehmensrechnung“ und „Einführung in das Controlling“ (10. Auflage, 650 Seiten). Wenn Sie sich für keine der Publikationen entscheiden, können Sie stattdessen wahlweise auch an der Verlosung von mehreren 8 GB iPods nano von Apple im Wert von je 249 Euro teilnehmen.

Wir versichern Ihnen, dass alle Angaben anonym und streng vertraulich behandelt werden. Sie können den Online-Fragebogen über den nachfolgenden Link aufrufen:

<http://www.unipark.de/uc/projektsteuerung/?code=XYZ>. Für Fragen und Anregungen stehen wir Ihnen telefonisch (0261/6509-482) und per E-Mail (projektmanagement@whu.edu) gerne zur Verfügung.

Wir danken Ihnen herzlich für Ihre Unterstützung!

Mit freundlichen Grüßen,

Prof. Dr. Dr. h. c. Jürgen Weber und Dipl.-Wi.-Ing. Matthias Mahlendorf

WHU – Otto Beisheim School of Management

Lehrstuhl für Controlling und Telekommunikation

– Stiftungslehrstuhl der Deutschen Telekom AG –

Wenn Sie sicher sind, dass Sie weder jetzt noch später an der Umfrage teilnehmen möchten, senden Sie uns bitte eine kurze Nachricht. Wir löschen Ihre Daten dann selbstverständlich umgehend aus unserer Befragung. Wir sind jedoch überzeugt, dass diese Studie sowohl für Sie persönlich als auch für Ihr Unternehmen und die Forschung hohen Nutzen stiftet, und würden uns deshalb sehr freuen, wenn Sie dem Link folgen und den Fragebogen ausfüllen.

<http://www.unipark.de/uc/projektsteuerung/?code=XYZ>

A.5 Fragebogen

Prof. Dr. Dr. h. c. Jürgen Weber
Lehrstuhl für Controlling und Telekommunikation
WHU – Otto Beisheim School of Management
Burgplatz 2, 56179 Vallendar



Erfolgreiche Steuerung risikoreicher Projekte *Die deutschlandweit größte Umfrage zum Projektcontrolling*



Prof. Dr. Dr. h. c. Jürgen Weber



Dipl.-Wi.-Ing. Matthias Mahlendorf

Von Ihrer Teilnahme hängt der Erfolg dieses Forschungsprojekts ab!

Nur durch die Auskünfte von Experten wie Ihnen können wir praxisbezogene Erkenntnisse gewinnen, die auch Ihnen zugute kommen.

Ihre Angaben werden anonym ausgewertet und streng vertraulich behandelt!

Die Ergebnisse werden ausschließlich in aggregierter Form veröffentlicht, so dass keine Rückschlüsse auf Sie persönlich oder Ihr Unternehmen möglich sind.

Das Ausfüllen wird insgesamt etwa 20 Minuten dauern.

Als Dankeschön für Ihre Unterstützung erhalten Sie nach Abschluss der Untersuchung die Auswertung der Ergebnisse in Form einer umfassenden Managementstudie, in der wir Ihnen wissenschaftlich fundierte, praxisorientierte Handlungsempfehlungen aufzeigen. Außerdem erhalten Sie eine Checkliste mit aktuellen Maßnahmen, mit denen Sie ihre Projektsteuerung zukünftig noch weiter verbessern können.

Zusätzlich können Sie sich als Dankeschön für eine der folgenden Prämien entscheiden, die Sie von uns kostenfrei erhalten:

- einen Buchband von Prof. Jürgen Weber, z. B. „Erfahrungen mit Value Based Management“, „Erfahrungen mit der Balanced Scorecard revisited“, „Investitionscontrolling in deutschen Großunternehmen“, „Strategisches Controlling“, „IFRS revisited: Quo vadis Unternehmensrechnung“ oder das Standardwerk „Einführung in das Controlling“ (10. Auflage, 650 Seiten),
- die kostenfreie Konferenzteilnahme für Sie oder einen Ihrer Mitarbeiter am „5. WHU Campus for Supply Chain Management“ Anfang 2008 im Wert von 260 Euro,
- die Teilnahme an der Verlosung von mehreren Apple iPod nano MP3-Playern (8 GB Speicher) im Wert von je 249 Euro.

Referenzen für unsere Forschungsprojekte und Hintergrundinformationen zu unserem Lehrstuhl finden Sie auf unserer Webseite <http://www.whu.edu/cms/index.php?id=2618>.

Für Rückfragen steht Ihnen Dipl.-Wi.-Ing. Matthias Mahlendorf gerne zur Verfügung:

Tel.: 0261/6509-482; Fax: 0261/6509-479

E-Mail: Projektmanagement@whu.edu

Zum Start des Fragebogens klicken Sie bitte auf „Weiter“.

Seite 1

Hatten Sie schon einmal mit einem Projekt zu tun, das abgebrochen wurde, bevor das eigentliche Projektziel erreicht wurde?

Ein paar Beispiele:

Ein Entwicklungsprojekt, bei dem irgendwann festgestellt wurde, dass das Ziel technisch nicht erreicht werden kann oder dass das Produkt nicht mehr benötigt wird.

Ein internes Projekt zur Optimierung oder Prozessverbesserung, das erfolglos eingestellt wurde.

Ein Strategiprojekt (z. B. Kauf eines Unternehmens oder Expansion in andere Märkte), das beendet wurde, ohne dass das ursprüngliche Ziel erreicht wurde.

etc.

Ja Nein

Hatten Sie schon einmal mit einem Projekt (im weitesten Sinne) zu tun, das gescheitert ist, d. h. zu Ende war, ohne dass das ursprüngliche Ziel erreicht wurde?

Ein paar Beispiele:

Die Einführung eines neuen Produkts, das am Markt jedoch nicht erfolgreich war.

Die Implementierung einer Software, die jedoch nie das Ziel erreichte, das vorgesehen war.

etc.

Ja Nein

Hatten Sie schon einmal mit einem Projekt (im weitesten Sinne) zu tun, bei dem ein Neustart/Relaunch durchgeführt wurde (z. B., um es wieder auf den richtigen Weg zu bringen)?

Ja Nein

Ihre Antworten auf die hier gestellten Fragen beeinflussen den weiteren Verlauf des Fragebogens. Bitte antworten Sie deshalb besonders sorgfältig.

Seite 2

Bitte erinnern Sie sich an abgebrochene Projekte (d. h. Projekte, die beendet wurden, bevor das Projektziel erreicht wurde), mit denen Sie in der Vergangenheit zu tun hatten. Wählen Sie jetzt gedanklich unter diesen Projekten dasjenige aus, an das Sie sich am besten erinnern können. (Dabei kann es sich auch um ein Projekt bei einem Ihrer früheren Arbeitgeber handeln.)

Wichtig: Beziehen Sie alle folgenden Antworten auf dieses eine bestimmte abgebrochene Projekt!

Um welche Art von Projekt handelte es sich bei dem abgebrochenen Projekt?

- Grundlagenforschung
- Entwicklung
- Implementierung (z. B. Software)
- Optimierung/Prozessverbesserung
- Outsourcing
- Produktneueinführung
- Markteintritt
- Anlagenbau
- Immobilienbau
- anderes: ...

Waren Sie bei dem Projekt derjenige, der hauptverantwortlich über Abbruch oder Fortsetzung des Projekts entscheiden musste? (Diese Person wird im Folgenden „Entscheider“ genannt.)

- Ja, ich war der hauptverantwortliche Entscheider.
- Nein, ich war NICHT der hauptverantwortliche Entscheider.
- Es gab keinen Hauptentscheider, sondern mehrere genau gleichrangige Entscheider – ich war einer von ihnen.
- Es gab keinen Hauptentscheider, sondern mehrere genau gleichrangige Entscheider – ich gehörte NICHT dazu.

Welche Stellung hatte der Entscheider in der Unternehmenshierarchie?

- Der Entscheider war gleichzeitig Inhaber des Unternehmens.
- Der Entscheider war nicht Inhaber, hatte aber eine Position auf der obersten Hierarchieebene inne (z. B. CEO).
- Der Entscheider befand sich nicht auf der obersten Hierarchieebene.

Bitte bewerten Sie die Qualität der folgenden Aspekte bei dem von Ihnen gewählten Projekt im Vergleich zu anderen abgebrochenen Projekten, die Sie kennen.	Viel schlechter			Genau gleich			Viel besser	Weiß nicht
	-3	-2	-1	0	1	2	3	
Zeitpunkt der Projektabbruchs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fundierung der Abbruchentscheidung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Qualität der Projektergebnisse (soweit sie fertiggestellt worden sind)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kosten (Budgeteinhaltung, effizienter Ressourceneinsatz etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zeit (Dauer, Pünktlichkeit etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Welcher Anteil der ursprünglich (bei Projektbeginn) geplanten Ressourcen wurde bis zum Abbruch des Projekts eingesetzt?

Angaben in Prozent der ursprünglich geplanten Gesamtmenge. Also z. B. 50 %, wenn bis zum Abbruch die Hälfte der zu Beginn geplanten Ressourcen investiert wurde, oder 200 %, wenn das Doppelte der ursprünglich geplanten Ressourcen investiert wurde.

Anteil der Manntage im Vergleich zur Planung ... %

Anteil Kosten (finanzielle Ressourcen) im Vergleich zur Planung ... %

Dauer des Projekts (Kalenderzeit) im Vergleich zur Planung ... %

Welcher Anteil der bei Projektbeginn geplanten Gesamtleistung wurde bis zum Abbruch des Projekts fertiggestellt?

Angabe bitte in Prozent

Grad der Fertigstellung ... %

	Trifft gar nicht zu			Trifft mittelmäßig zu			Trifft voll zu	Weiß nicht
	0	1	2	3	4	5	6	
Das Projekt war sehr komplex.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Projekt war sehr risikoreich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Projekt war stark abhängig von äußeren Faktoren (z. B. Kunden, Gesetzgeber etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte bewerten Sie die folgenden Aussagen über den Zeitpunkt der ersten Probleme des Projekts.	Trifft gar nicht zu			Trifft mittelmäßig zu			Trifft voll zu	Weiß nicht
	0	1	2	3	4	5	6	
Die Probleme sind erst lange nach Projektbeginn aufgetreten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Gefahren für das Projekt kamen erst sehr spät zum Vorschein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die ersten auffälligen Schwierigkeiten tauchten erst lange Zeit nach dem Projektstart auf.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte schätzen Sie, wie viele Ressourcen, absolut betrachtet, insgesamt bis zum Abbruch in das Projekt geflossen sind.

Personeneinsatz – Angabe bitte in Anzahl der Manntage ...

Kosten (finanzielle Ressourcen) – Angabe bitte in TAUSEND Euro ...

Dauer des Projekts (Kalenderzeit, die von Beginn bis Ende des Projekts vergangen ist) – Angabe bitte in Anzahl der Kalendermonate ...

	Nie		Manchmal		Häufig		Immer	Weiß nicht
	0	1	2	3	4	5	6	
Ich war sehr überzeugt, dass die Schätzungen der zukünftigen Entwicklung korrekt seien.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich war mir sicher in Bezug auf die Prognosen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich war zuversichtlich, dass sich die Schätzungen für den Erfolg des Projekts bestätigen würden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hatte mich fest auf die Zukunftsprognosen verlassen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die folgenden Fragen sind aus wissenschafts-methodischen Gründen notwendig. Bitte beantworten Sie auch diese Fragen so präzise wie möglich, selbst wenn die Fragen teilweise nicht zum eigentlichen Thema zu gehören scheinen. Bitte geben Sie an, inwieweit die folgenden Aussagen auf Sie zutreffen.

	Trifft gar nicht zu			Trifft mittelmäßig zu			Trifft voll zu	Weiß nicht
	0	1	2	3	4	5	6	
Manchmal erzähle ich Lügen, wenn ich muss.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich vertusche meine Fehler nie.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe schon Schlechtes über einen Freund hinter seinem Rücken gesagt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich nehme nie Dinge, die mir nicht gehören.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich tratsche nicht über anderer Leute Angelegenheiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Nur für Nicht-Entscheider]	Trifft gar nicht zu			Trifft mittelmäßig zu			Trifft voll zu	Weiß nicht
	0	1	2	3	4	5	6	
Ich habe ein gutes persönliches Verhältnis zu dem Entscheider.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich mit dem Entscheider freundschaftlich verbunden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich verstehe mich gut mit dem Entscheider.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Seite 6

	Trifft gar nicht zu			Trifft mittelmäßig zu			Trifft voll zu	Weiß nicht
	0	1	2	3	4	5	6	
Ich habe umfassende Kenntnisse über das Projekt, zu dem ich diesen Fragebogen ausgefüllt habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In welcher Funktion waren Sie während des damaligen Projekts tätig?

- Human resources
- Controlling/Finanzen
- Einkauf
- Berater/Consultant
- Produktion
- Supply Chain Management/Logistik
- Forschung & Entwicklung
- IT
- General Management (funktionsübergreifend)
- andere

	Viel zu kurz			Genau richtig			Viel zu lang
	-3	-2	-1	0	1	2	3
Für das untersuchte Thema finde ich die Länge des Fragebogens ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte füllen Sie die folgenden Felder aus, damit wir Sie bezüglich des Dankeschöns kontaktieren können.

Die mit * markierten Angaben sind notwendig, wenn wir Ihnen ein Dankeschön per Post zuschicken sollen. Die E-Mail-Adresse benötigen wir, um Ihnen den Best-Practice-Bericht senden zu können.

Vorname* :

Nachname* :

Position :

E-Mail-Adresse* :

Unternehmen :

Straße* :

Postleitzahl* :

Stadt* :

Telefonnummer :

Bitte wählen Sie nun Ihr „Dankeschön“ aus.

In jedem Fall erhalten Sie die Ergebnisse der Haupt-Umfrage in Form eines Best-Practice-Berichts und die Checkliste mit aktuellen Maßnahmen zur Verbesserung der Projektsteuerung. Zusätzlich können Sie unter folgenden Optionen wählen:

- Ich benötige keine der Publikationen und möchte auch weder an der Konferenz noch an der iPod-Verlosung teilnehmen.
- Ich möchte an der Verlosung teilnehmen mit der Chance von 1 zu 100, einen 8 GB iPod nano von Apple im Wert von 249 Euro zu gewinnen.
- Ich (oder ein Mitarbeiter von mir) möchte kostenlos an der Konferenz „5. WHU Campus for Supply Chain Management“ Anfang 2008 im Wert von 260 Euro teilnehmen.
- Ich möchte eine kostenlose Publikation von Prof. Dr. Dr. h. c. Jürgen Weber (Auswahl des Titels siehe nächste Frage).

Bitte wählen Sie nun eine der folgenden Publikationen.

Die Auswahl gilt nur, wenn Sie oben als Dankeschön eine Publikation ausgewählt haben, nicht hingegen, wenn Sie stattdessen die Teilnahme an der Verlosung des iPods oder die Konferenzteilnahme angekreuzt haben.

- Investitionscontrolling in deutschen Großunternehmen – Ergebnisse einer Benchmarking-Studie
- IFRS revisited – Quo vadis Unternehmensrechnung?
- Erfahrungen mit der Balanced Scorecard Revisited
- Strategisches Controlling
- Einführung in das Controlling (10. Auflage)
- Erfahrungen mit Value Based Management
- Was machen Controller wann warum?

Wenn Sie Anmerkungen oder Kommentare zu dieser Umfrage haben, können Sie diese gern hier eintragen. Besonders wichtig ist es für uns zu wissen, wenn bestimmte Fragen für Ihre Situation nicht gepasst haben.

...

Abschicken des Fragebogens:

Mit einem Klick auf das Feld „Weiter“ bestätigen Sie Ihre Eingaben und senden den Fragebogen ab.

Sie können danach keine Angaben mehr korrigieren und nicht auf vorherige Seiten zurückkehren.

Wenn Sie den Fragebogen erst vollständig ausfüllen möchten, klicken Sie hier nicht „Weiter“ sondern auf „Zurück“, bis Sie die noch unausgefüllten Fragen erreichen.

A.6 Erstes Erinnerungsschreiben

Sehr geehrte/r Frau/Herr XYZ,

Entscheidungen über risikoreiche Projekte stellen eine der größten Herausforderungen für das technische Management dar. Zu diesem sicherlich auch für Sie relevanten Thema führt unser Lehrstuhl derzeit ein umfassendes Forschungsprojekt durch. Ein Teil des Forschungsprojekts „Erfolgreiche Steuerung risikoreicher Projekte“ besteht in der Befragung von technischen Leitern aus verschiedenen Branchen. Wir möchten Sie bitten, durch Ausfüllen eines Online-Fragebogens zum Erfolg des Forschungsprojekts beizutragen.

Als Dankeschön für das Ausfüllen des Fragebogens (Dauer: ca. 20 Min.) erhalten Sie nach Auswertung der Ergebnisse unter anderem einen exklusiv erstellten Best-Practice-Bericht. Basierend auf den Ergebnissen dieses Forschungsprojekts werden praxisorientierte Handlungsempfehlungen für Ihr Unternehmen vorgestellt. Außerdem erhalten Sie eine Checkliste mit aktuellen Maßnahmen zur Verbesserung der Projektsteuerung.

Zusätzlich können Sie sich eine der folgenden Prämien aussuchen, die wir Ihnen kostenlos als Dankeschön anbieten:

- einen Buchband von Prof. Jürgen Weber; z. B. „Erfahrungen mit der Balanced Scorecard revisited“, „Erfahrungen mit Value Based Management“, „Investitionscontrolling in deutschen Großunternehmen“, „Strategisches Controlling“, „IFRS revisited: Quo vadis Unternehmensrechnung“ oder das Standardwerk „Einführung in das Controlling“ (10. Auflage, 650 Seiten),
- die kostenfreie Teilnahme an der Konferenz „5. WHU Campus for Supply Chain Management“ Anfang 2008 im Wert von 260 Euro für Sie oder einen Ihrer Mitarbeiter,
- die Teilnahme an der Verlosung von mehreren Apple iPod nano MP3-Playern (8 GB Speicher) im Wert von je 249 Euro.

Wir versichern Ihnen, dass alle Angaben anonym und streng vertraulich behandelt werden.

Sie können den Online-Fragebogen über den nachfolgenden Link aufrufen:

<http://www.unipark.de/uc/projektsteuerung1/?code=XYZ>.

Für Fragen und Anregungen stehen wir Ihnen telefonisch (0261/6509-482) und per E-Mail (projektsteuerung@whu.edu) gerne zur Verfügung.

Wir danken Ihnen herzlich für Ihre Unterstützung!

Mit freundlichen Grüßen,

Prof. Dr. Dr. h. c. Jürgen Weber und Dipl.-Wi.-Ing. Matthias Mahlendorf

WHU – Otto Beisheim School of Management

Lehrstuhl für Controlling und Telekommunikation

– Stiftungslehrstuhl der Deutschen Telekom AG –

PS: Wenn Sie noch unschlüssig sind, ob diese Studie tatsächlich relevante Ergebnisse für Sie zeigen wird, empfehlen wir Ihnen einen Blick auf die Referenzunternehmen, die bereits an unseren Forschungsprojekten teilgenommen haben:

<http://www.whu.edu/cms/index.php?id=2618>.

PPS: Wenn Sie sicher sind, dass Sie weder jetzt noch später an der Umfrage teilnehmen möchten, senden Sie uns bitte eine kurze Nachricht. Wir löschen Ihre Daten dann selbstverständlich umgehend aus unserer Befragung. Wir sind jedoch überzeugt, dass diese Studie sowohl für Sie persönlich als auch für Ihr Unternehmen und die Forschung hohen Nutzen stiftet, und würden uns deshalb sehr freuen, wenn Sie dem Link folgen und den Fragebogen ausfüllen.

<http://www.unipark.de/uc/projektsteuerung1/?code=XYZ>

A.7 Abschließendes Erinnerungsschreiben

Sehr geehrte/r Frau/Herr XYZ,

was tun Sie, damit Projekte nicht aus dem Ruder laufen und Entscheidungen möglichst objektiv getroffen werden?

Wie kann verhindert werden, dass kleine menschliche Schwächen zu folgenschweren Fehlentscheidungen für Ihr Unternehmen führen?

Da Controllinginstrumente bislang die menschlichen Aspekte häufig vernachlässigen, führt unser Lehrstuhl für Controlling gegenwärtig eine deutschlandweite Umfrage durch, mit deren Hilfe Antworten auf diese und weitere Fragen gegeben werden sollen. Unsere Umfrage ist von den angeschriebenen Controllern, Projektmanagern und technischen Leitern erfreulicherweise sehr positiv aufgenommen worden, wie die sehr hohe Beteiligung von bislang 885 Personen zeigt. Die Online-Befragung wird am 10. August beendet, so dass bis dahin die letzte Chance besteht, sich an der Umfrage zu beteiligen und die Ergebnisse der Studie, in Form einer Managementstudie mit konkreten Praxisempfehlungen, zu erhalten.

Wenn Sie an der Umfrage teilnehmen möchten, klicken Sie bitte einfach auf den folgenden Link: <http://www.unipark.de/uc/projektcontrolling1/?code=XYZ>.

Falls Sie schon mit dem Ausfüllen begonnen haben, sind Ihre bisherigen Angaben im System hinterlegt und Sie brauchen nur noch die fehlenden Fragen zu beantworten. Entsprechend werden Sie vom System direkt auf die Seite geleitet, auf der Sie beim letzten Mal die Beantwortung unterbrechen mussten.

Für Rückfragen stehen wir Ihnen gern zur Verfügung.

Mit freundlichen Grüßen,

Prof. Dr. Dr. h. c. Jürgen Weber und Dipl.-Wi.-Ing. Matthias Mahlendorf

WHU – Otto Beisheim School of Management
Lehrstuhl für Controlling und Telekommunikation
– Stiftungslehrstuhl der Deutschen Telekom AG –
Tel.: 0261-6509-471
E-Mail: Projektcontrolling@whu.edu

PS: Wenn Sie nicht teilnehmen möchten, brauchen Sie nichts weiter zu tun. Wir werden Sie nicht mehr anschreiben.

PPS: Als Dankeschön können Sie sich zusätzlich zur Auswertung der Daten in Form eines Best-Practice-Berichts und einer Checkliste mit aktuellen Maßnahmen im Projektcontrolling eine der folgenden Prämien aussuchen, die Sie von uns kostenlos erhalten:

- einen Buchband aus der Reihe Advanced Controlling (z.B. „Erfahrungen mit der Balanced Scorecard revisited“ oder „IFRS revisited: Quo vadis Unternehmensrechnung“)
- oder die Teilnahme an der Verlosung von mehreren Apple iPod nano MP3-Playern (8 GB Speicher) im Wert von je 249 Euro.

<http://www.unipark.de/uc/projektcontrolling1/?code=XYZ>

A.8 Mittelwerte und Standardabweichungen

Tabelle A.2: Latente Konstrukte

Konstruktbezeichnung (Ankerpunkte der Skala)		M	SD
Nr.	Wortlaut des Indikators		
Verspätung des Projektabbruchs (0 = trifft gar nicht zu ... 6 = trifft voll zu)		3,2	2,0
Im Nachhinein betrachtet, ...			
1.	... wurde das Projekt zu spät abgebrochen.	3,1	2,2
2.	... wurde zu hartnäckig an dem Projekt festgehalten.	3,2	2,1
3.	... hätte das Projekt früher abgebrochen werden müssen.	3,3	2,1
Beurteilung des Ressourcen-Einsatzes (-3 = viel zu wenig ... +3 = viel zu viel)		0,5	1,3
Bitte bewerten Sie, ob zu wenig, genau richtig oder zu viele Ressourcen in das abgebrochene Projekt geflossen sind.			
1.	Mannmonate	0,2	1,6
2.	Kosten (finanzielle Ressourcen)	0,4	1,4
3.	Dauer des Projekts (Kalenderzeit)	0,8	1,5
investierter Anteil der geplanten Ressourcen (Angabe in Prozent)		94,8	61,0
Welcher Anteil der ursprünglich (bei Projektbeginn) geplanten Ressourcen wurde bis zum Abbruch des Projekts eingesetzt?			
1.	Mannmonate ... %	89,9	63,0
2.	Kosten (finanzielle Ressourcen) ... %	84,5	61,8
3.	Dauer des Projekts (Kalenderzeit) ... %	110,0	85,5
Wahrnehmungsschwelle (0 = nie ... 6 = immer)		2,1	1,2
Wichtige negative Informationen ...			
1.	... sind für mich in der Menge der Informationen untergegangen.		eliminiert
2.	... waren für mich schwer zu bemerken.		eliminiert
3.	... waren für mich schwer von den unwichtigen Informationen zu unterscheiden.	2,0	1,3
4.	... waren zu unauffällig, um von mir gut bemerkt zu werden.	2,0	1,3
5.	... waren schwer durch mich als solche zu identifizieren.	2,2	1,4
selektive Wahrnehmung (0 = nie ... 6 = immer)		2,7	1,4
Wenn neue Informationen zum Projekt auftraten, ...			
1.	... schenkte ich positiven Informationen größeres Interesse als negativen.		eliminiert
2.	... habe ich Informationen weniger Beachtung geschenkt, wenn sie meinen Vorstellungen widersprachen.	2,6	1,6
3.	... wandte ich nur bestimmten Informationen meine Aufmerksamkeit zu (d. h. unterlag selektiver Wahrnehmung).	2,7	1,5
4.	... fokussierte sich meine Aufmerksamkeit auf meine „Lieblingsthemen“.	2,7	1,7

Die Items betreffen das gesamte Sample (N = 677), wobei Nicht-Entscheidungsverantwortliche eine Version des Fragebogens erhielten, in der das Wort „ich“ durch „der Entscheider“ ersetzt war.

M = arithmetischer Mittelwert, SD = Standardabweichung

Konstruktbezeichnungen sowie M und SD für Konstrukte sind fett hervorgehoben.

Tabelle A.2: Latente Konstrukte (Fortsetzung von Seite 236)

Konstruktbezeichnung (Ankerpunkte der Skala)		M	SD
Nr.	Wortlaut des Indikators		
Self-Justification-Druck (0 = nie ... 6 = immer)		2,2	1,6
1.	Ich hatte Sorge um mein Image, falls es zu einem Projektabbruch kommen würde.	2,1	1,8
2.	Ich befürchtete, dass wichtige Personen (z. B. Vorgesetzte, Kapitalgeber u. Ä.) bei einem Projektabbruch einen schlechten Eindruck von mir haben könnten.	2,1	1,8
3.	Ich stand unter hohem Druck, das Projekt erfolgreich abzuschließen.	eliminiert	
4.	Ich empfand den Erfolg des Projekts als eng verknüpft mit meinem eigenen, persönlichen Erfolg.	2,5	1,9
Sunk-Cost-Effekt (0 = nie ... 6 = immer)		2,4	1,5
1.	Einen Projektabbruch sah ich wegen der bereits investierten unwiederbringlichen Investitionen (Sunk Costs) als Verschwendung von Ressourcen an.	2,3	1,7
2.	Ich sprach mich für eine Projektfortführung aus, weil bei einem Projektabbruch die zuvor erfolgten Investitionen umsonst gewesen wären.	2,5	1,8
3.	Einen Projektabbruch habe ich als Vernichtung von bereits investierten Ressourcen betrachtet.	2,5	1,8
4.	Ein Abbruch des Projekts erschien mir zunehmend teurer, je mehr Ressourcen in das Projekt geflossen waren.	2,7	1,7
Optimismus (0 = nie ... 6 = immer)		3,9	1,3
1.	Ich bewertete das Projekt lange Zeit zu positiv.	eliminiert	
2.	Ich war lange Zeit optimistisch, was den erwarteten Ausgang des Projekts betraf.	3,9	1,4
3.	Auch wenn es Probleme gab, war ich zuversichtlich, dass das Projekt am Ende erfolgreich sein würde.	3,9	1,4
4.	Ich war überzeugt vom erfolgreichen Abschluss des Projekts.	4,0	1,4
übermäßiger Optimismus (0 = trifft gar nicht zu ... 6 = trifft voll zu)			
1.	Ich bewertete das Projekt lange Zeit zu positiv.	3,2	1,6
Zeitpunkt des ersten Problemauftretens		2,8	1,7
1.	Die Probleme sind erst lange nach Projektbeginn aufgetreten.	2,8	1,8
2.	Die Gefahren für das Projekt kamen erst sehr spät zum Vorschein.	2,9	1,8
3.	Die ersten auffälligen Schwierigkeiten tauchten erst lange Zeit nach dem Projektstart auf.	2,8	1,9
nicht hinterfragter Freiraum (0 = nie ... 6 = immer)		3,1	1,5
Bei den Entscheidungen über Abbruch oder Fortsetzung des Projekts ...			
1.	... hatte ich einen großen Freiraum, eigene Interessen einfließen zu lassen.	3,6	1,7
2.	... musste ich mich nur vor mir selbst rechtfertigen.	2,8	2,1
3.	... wurden meine Entscheidungen ohne Hinterfragen akzeptiert.	3,1	1,8
4.	... blieb es mir überlassen, ob ich meine Gründe offenlegen wollte.	2,9	2,0

Die Items betreffen das gesamte Sample (N = 677), wobei Nicht-Entscheidungsverantwortliche eine Version des Fragebogens erhielten, in der das Wort „ich“ durch „der Entscheider“ ersetzt war.

M = arithmetischer Mittelwert, SD = Standardabweichung

Konstruktbezeichnungen sowie M und SD für Konstrukte sind fett hervorgehoben.

Tabelle A.2: Latente Konstrukte (Fortsetzung von Seite 237)

Konstruktbezeichnung (Ankerpunkte der Skala)		M	SD
Nr.	Wortlaut des Indikators		
Overconfidence (0 = nie ... 6 = immer)		4,0	1,1
1.	Ich war sehr überzeugt, dass die Schätzungen der zukünftigen Entwicklung korrekt seien.	4,0	1,2
2.	Ich war mir sicher in Bezug auf die Prognosen.	3,9	1,2
3.	Ich war zuversichtlich, dass sich die Schätzungen für den Erfolg des Projekts bestätigen würden.	3,9	1,3
4.	Ich hatte mich fest auf die Zukunftsprognosen verlassen.	eliminiert	
Kapitalmarktorientierung (0 = trifft gar nicht zu ... 6 = trifft voll zu)		1,8	1,8
1.	Das Unternehmen war auf den Kapitalmarkt angewiesen.	1,6	2,0
2.	Das Unternehmen bemühte sich um eine gute Kommunikation mit Kapitalmarktakteuren (z. B. Aktionären, Analysten, Gläubigern).	2,1	2,1
3.	Kapitalmarktorientierung war im Unternehmen groß geschrieben.	1,8	2,0
selektive Wahrnehmung anderer Manager (0 = nie ... 6 = immer)		2,2	1,2
Ein anderer, durchschnittlich guter Manager hätte damals in meiner Situation ...			
1.	... den positiven Informationen mehr Interesse geschenkt als den negativen.	eliminiert	
2.	... Informationen weniger Beachtung geschenkt, wenn sie seinen Vorstellungen widersprochen hätten.	2,2	1,3
3.	... nur bestimmten Informationen seine Aufmerksamkeit zugewandt (d. h., wäre selektiver Wahrnehmung unterlegen).	2,2	1,3
4.	... seine Aufmerksamkeit auf seine „Lieblingsthemen“ fokussiert.	2,3	1,3
Freundschaft* (0 = trifft gar nicht zu ... 6 = trifft voll zu)		3,2	1,3
1.	Ich habe ein gutes persönliches Verhältnis zu dem Entscheider.	3,5	1,6
2.	Ich fühle mich mit dem Entscheider freundschaftlich verbunden.	2,2	1,7
3.	Ich verstehe mich gut mit dem Entscheider.	3,8	1,5

* Die Items des Konstrukts Freundschaft betreffen nur die Nicht-Entscheidungsverantwortlichen (N = 390). Alle übrigen Items betreffen das gesamte Sample (N = 677), wobei Nicht-Entscheidungsverantwortliche eine Version des Fragebogens erhielten, in der das Wort „ich“ durch „der Entscheider“ ersetzt war.

M = arithmetischer Mittelwert, SD = Standardabweichung

Konstruktbezeichnungen sowie M und SD für Konstrukte sind fett hervorgehoben.

Tabelle A.3: Maßnahmen

Maßnahmengruppe (Ankerpunkte der Skala)			
Nr.	Wortlaut des Indikators	M	SD
Maßnahmen zur Reduzierung der Wahrnehmungsschwelle (0 = nie ... 6 = immer)			
Meilensteine	Es wurden Meilensteine festgelegt, d. h. messbare Zwischenziele mit Deadline (Termin).	3,6	1,8
Genehmigung	Das Projekt musste jeweils vor Beginn der nächsten Teilphase eine Genehmigung erhalten, bevor es fortgesetzt wurde.	2,3	2,0
aktive Entscheidung	Es wurde regelmäßig (z. B. an jedem Meilenstein) aktiv über die Fortführung des Projekts entschieden.	3,2	1,9
Berichte	Es wurden Projektfortschrittsberichte (z. B. mit Kennzahlen über Abweichungen bei Kosten, Zeit, Qualität) für mich erstellt.	3,1	1,9
eindeutige Kriterien	Es gab eindeutige Beurteilungskriterien, nach denen die Projektinformationen bewertet werden konnten (z. B. Zielwerte oder Benchmarkingstudien).	2,9	1,8
Quantifizierung	Die Informationen über den Projektverlauf wurden für mich quantifiziert (d. h., mir lagen konkrete Zahlenwerte vor).	3,5	1,7
Maßnahmen zur Reduzierung der selektiven Wahrnehmung (0 = nie ... 6 = immer)			
Gefahrenhinweis	Ich wurde explizit auf die Gefahr hingewiesen, dass Projekte häufig zu spät abgebrochen werden (z. B. anhand von Daten vergangener Projekte).	1,8	1,7
Abbruchhinweis	Mir wurde verdeutlicht, dass ich das Projekt abbrechen muss, wenn die Ziele nicht mehr erreicht werden können.	2,9	2,0
Maßnahmen zur Reduzierung des Self-Justification-Drucks (0 = nie ... 6 = immer)			
Mitverantwortung	Es gab andere Personen (z. B. Controller), die explizit Mitverantwortung für das Fehlschlagen des Projekts übernommen haben.	2,3	2,9
Prozessqualität	Ich wurde nicht nur anhand der tatsächlichen Folgen meiner Entscheidungen bewertet, sondern auch für die Qualität meiner Entscheidungsprozesse (d. h., es wurde auch beurteilt, ob die Entscheidung zum jeweiligen Zeitpunkt gut begründet war).	3,0	1,9
Ansehen	Mir wurde klar kommuniziert, dass mein Ansehen bei einem Projektabbruch keinerlei Schaden nehmen würde.	2,9	2,2

Die Items betreffen das gesamte Sample (N = 677), wobei Nicht-Entscheidungsverantwortliche eine Version des Fragebogens erhielten, in der das Wort „ich“ durch „der Entscheider“ ersetzt war.

M = arithmetischer Mittelwert, SD = Standardabweichung

Tabelle A.3: Maßnahmen (Fortsetzung von Seite 239)

Maßnahmengruppe (Ankerpunkte der Skala)			
Nr.	Wortlaut des Indikators	M	SD
Maßnahmen zur Reduzierung des Sunk-Cost-Effekts (0 = nie ... 6 = immer)			
expliziter Ausweis	Die Sunk Costs (d. h. in der Vergangenheit investierte Kosten, die nicht zurückgewonnen werden können) wurden ausdrücklich als solche ausgewiesen.	2,4	2,2
Irrelevanz gekennzeichnet	Die Sunk Costs wurden als irrelevant für Entscheidungen gekennzeichnet (d. h., es war deutlich, dass nur zukünftige Kosten und Erlöse entscheidungsrelevant sind).	3,3	2,0
Verfahren	Mir standen angemessene Verfahren für die Bewertung des Projekts zur Verfügung.	3,4	1,8
Expertenunterstützung	Ich wurde bei der Anwendung der Bewertungsinstrumente durch einen Experten (z. B. Controller) unterstützt.	3,0	2,2
Alternativen	In der Projektberichterstattung wurde auf die Möglichkeit hingewiesen, die Ressourcen für das scheiternde Projekt in andere Projekte zu investieren.	2,2	1,9
Maßnahmen zur Reduzierung des (übermäßigen) Optimismus (0 = nie ... 6 = immer)			
Analogien	Schätzungen über den voraussichtlichen Projekterfolg wurden mit Hilfe von Analogien/Vergleichen zu anderen Projekten durchgeführt.	2,5	1,8
Advocatus Diaboli	Es gab jemanden, der die Rolle des „Advocatus Diaboli“ übernommen hat (d. h., jemand hat bewusst pessimistisch argumentiert).	2,3	1,9
Maßnahmen zur Reduzierung des nicht hinterfragten Freiraums (0 = nie ... 6 = immer)			
Zielbekanntgabe	Die Ziele, die ich mit der Weiterführung des Projekts verfolgt habe, wurden bekanntgegeben.	4,0	1,8
Zieldiskussion	Meine Projektziele wurden diskutiert (z. B. in Bezug auf ihre Erreichbarkeit).	3,6	1,7
Abbruchkriterien	Bereits zu Beginn des Projekts (und z. B. an Meilensteinen) wurden Abbruchkriterien festgelegt, die den Projektbeteiligten bekannt waren (d. h. Zwischenziele, bei deren Verfehlen das Projekt abgebrochen werden sollte).	1,8	1,9

Die Items betreffen das gesamte Sample (N = 677), wobei Nicht-Entscheidungsverantwortliche eine Version des Fragebogens erhielten, in der das Wort „ich“ durch „der Entscheider“ ersetzt war.

M = arithmetischer Mittelwert, SD = Standardabweichung

Literaturverzeichnis

- Abdi**, Hervé (2007), Bonferroni and Sidak Corrections for Multiple Comparisons, in: N. J. Salkind (Hrsg.), *Encyclopedia of Measurement and Statistics*, Thousand Oaks, CA: Sage, S. 103–107.
- Abelson**, Robert P. und Andrew **Levi** (1985), Decision Making and Decision Theory, in: G. Lindzey und E. Aronson (Hrsg.), *The Handbook of Social Psychology*, 3. Aufl., Bd. 2, Hillsdale, NJ: Erlbaum, S. 231–309.
- Albers**, Sönke und Oliver **Götz** (2006), Messmodelle mit Konstrukten zweiter Ordnung in der betriebswirtschaftlichen Forschung, in: *Die Betriebswirtschaft*, 66. Jg., Nr. 6, S. 669–677.
- Albers**, Sönke und Lutz **Hildebrandt** (2006), Methodische Probleme bei der Erfolgsfaktorenforschung: Messfehler, formative versus reflektive Indikatoren und die Wahl des Strukturgleichungsmodells, in: *Zeitschrift für betriebswirtschaftliche Forschung*, 58. Jg., S. 2–33.
- Allen**, J. M. und K. P. **Norris** (1970), Project Estimates and Outcomes in Electricity Generating Research, in: *Journal of Management Studies*, 7. Jg., Nr. 3, S. 271–287.
- Allison**, Paul D. (2002), Missing Data, Bd. 07-136 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.
- American Psychological Association** (1977), *Standards for Educational & Psychological Tests*, Washington, DC: American Psychological Association.
- Anastasi**, Anne (1986), Evolving Concepts of Test Validation, in: *Annual Review of Psychology*, 37. Jg., S. 1–15.
- Anderson**, James C., David W. **Gerbing** und John E. **Hunter** (1987), On the Assessment of Unidimensional Measurement: Internal and External Consistency, and Overall Consistency Criteria, in: *Journal of Marketing Research*, 24. Jg., S. 432–437.
- Aquilino**, William S. (1994), Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment, in: *Public Opinion Quarterly*, 58. Jg., S. 210–240.
- Arbuckle**, James L. (2006), *Amos 7.0 User's Guide*, Chicago, IL: SPSS, Inc.
- Archer**, Robert P. (1997), *MMPI-A: Assessing Adolescent Psychopathology*, 2. Aufl., Mahwah, NJ: Erlbaum.
- Archer**, Robert P. und David E. **Elkins** (1999), Identification of Random Responding on the MMPI-A, in: *Journal of Personality Assessment*, 73. Jg., Nr. 3, S. 407–421.

- Arkes**, Hal R. und Catherine **Blumer** (1985), The Psychology of Sunk Cost, in: *Organizational Behavior and Human Decision Processes*, 35. Jg., S. 124–140.
- Armstrong**, J. Scott und Terry S. **Overton** (1977), Estimating Nonresponse Bias in Mail Surveys, in: *Journal of Marketing Research*, 14. Jg., Nr. 3, S. 396–402.
- Armstrong**, Robert W., Robert J. **Williams** und J. Douglas **Barrett** (2004), The Impact of Banality, Risky Shift and Escalating Commitment on Ethical Decision Making, in: *Journal of Business Ethics*, 53. Jg., S. 365–370.
- Ashley**, David B., Clive S. **Lurie** und Edward J. **Jaselskis** (1987), Determinants of Construction Project Success, in: *Project Management Journal*, 18. Jg., Nr. 2, S. 69–79.
- Backhaus**, Klaus, Boris **Blechsmidt** und Maik **Eisenbeiß** (2006), Der Stichprobeneinfluss bei Kausalanalysen, in: *Die Betriebswirtschaft*, 66. Jg., S. 711–726.
- Backhaus**, Klaus, Bernd **Erichson**, Wulff **Plinke** und Rolf **Weiber** (2006), *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, 11. Aufl., Berlin: Springer.
- Bagozzi**, Richard P. und Youjae **Yi** (1988), On the Evaluation of Structural Equation Models, in: *Journal of the Academy of Marketing Sciences*, 16. Jg., Nr. 1, S. 74–94.
- Bagozzi**, Richard P. und Youjae **Yi** (1991), Multitrait-Multimethod Matrices in Consumer Research, in: *Journal of Consumer Research*, 17. Jg., S. 421–458.
- Balderjahn**, Ingo (1986), *Das umweltbewußte Konsumentenverhalten: Eine empirische Studie*, Bd. 123 der Reihe *Betriebswirtschaftliche Schriften*, Berlin: Dunker & Humblot.
- Balderjahn**, Ingo (2008), Editorial: Totgeglaubte leben länger! Auch in der BWL?, in: *Die Betriebswirtschaft*, 3. Jg., S. 267–269.
- Bamber**, E. Michael (1993), Opportunities in Behavioral Accounting Research, in: *Behavioral Research in Accounting*, 5. Jg., S. 1–29.
- Barton**, Sidney L., Dennis **Duchon** und Kenneth J. **Dunegan** (1989), An Empirical Test of Staw and Ross's Prescriptions for the Management of Escalation of Commitment Behavior in Organizations, in: *Decision Sciences*, 20. Jg., Nr. 3, S. 532–544.
- Bateman**, Thomas S. (1986), The Escalation of Commitment and Sequential Decision-Making, in: *Decision Sciences*, 17. Jg., S. 33–49.
- Bazerman**, Max H. (1993), Negotiator Judgment: A Critical Look at the Rationality Assumptions, in: *American Behavioral Scientist*, 27. Jg., Nr. 2, S. 211–228.
- Bazerman**, Max H., Rafik I. **Beekun** und F. David **Schoorman** (1982), Performance Evaluation in a Dynamic Context: A Laboratory Study of the Impact of a Prior Commitment to the Ratee, in: *Journal of Applied Psychology*, 67. Jg., Nr. 6, S. 873–876.
- Beach**, Daniel A. (1989), Identifying the Random Responder, in: *Journal of Psychology*, 123. Jg., Nr. 1, S. 101–103.
- Bentler**, Peter M. (1990), Comparative Fit Indexes in Structural Models, in: *Psychological Bulletin*, 107. Jg., S. 238–246.
- Bentler**, Peter M. und Chih-Ping **Chou** (1987), Practical Issues in Structural Equation Modeling, in: *Sociological Methods and Research*, 16. Jg., S. 78–136.

- Bernaards**, Coen A. und Klaas **Sijtsma** (1999), Factor Analysis of Multidimensional Polytomous Item Response Data Suffering From Ignorable Item Nonresponse, in: *Multivariate Behavioral Research*, 34. Jg., S. 277–313.
- Bernardi**, Richard A. (2006), Associations between Hofstede's Cultural Constructs and Social Desirability Response Bias, in: *Journal of Business Ethics*, 65. Jg., S. 43–53.
- Bird**, Graham (2005), Over-optimism and the IMF, in: *The World Economy*, 28. Jg., Nr. 9, S. 1355–1373.
- Birnberg**, Jacob G. (1973), Empirical Research in Behavioral Accounting, in: *Management International Review*, 13. Jg., Nr. 2-3, S. 71–77.
- Birnberg**, Jacob G. (2000), The Role of Behavioral Research in Management Accounting Education in the 21st Century, in: *Issues in Accounting Education*, 15. Jg., Nr. 4, S. 713–728.
- Birnberg**, Jacob G. und Raghu **Nath** (1967), Implications of Behavioral Science for Managerial Accounting, in: *The Accounting Review*, 42. Jg., Nr. 3, S. 468–479.
- Birnberg**, Jacob G. und Jeffrey F. **Shields** (1989), Three Decades of Behavioral Accounting Research: A Search For Order, in: *Behavioral Research in Accounting*, 1. Jg., S. 23–74.
- Bobocel**, D. Ramona und John P. **Meyer** (1994), Escalating Commitment to a Failing Course of Action: Separating the Roles of Choice and Justification, in: *Journal of Applied Psychology*, 79. Jg., Nr. 3, S. 360–363.
- Bohm**, Peter (1972), Estimating Willingness to Pay: An Experiment, in: *European Economic Review*, 3. Jg., S. 111–130.
- Bolton**, Ruth N. (1993), Pretesting Questionnaires: Content Analyses of Respondents' Concurrent Verbal Protocols, in: *Marketing Science*, 12. Jg., Nr. 3, S. 280–303.
- Booz-Allen-Hamilton-Inc.** (Hrsg.) (1968), *Management of New Products*, 4. Aufl., New York, NY: Booz Allen Hamilton.
- Bosnjak**, Michael, Tracy L. **Tuten** und Werner W. **Wittmann** (2005), Unit (Non)Response in Web-Based Access Panel Surveys: An Extended Planned-Behavior Approach, in: *Psychology & Marketing*, 22. Jg., Nr. 6, S. 489–505.
- Boulding**, William, Ruskin **Morgan** und Richard **Staelin** (1997), Pulling the Plug to Stop the New Product Drain, in: *Journal of Marketing Research*, 34. Jg., S. 164–176.
- Bowen**, Michael G. (1987), The Escalation Phenomenon Reconsidered: Decision Dilemma or Decision Errors?, in: *Academy of Management Review*, 12. Jg., Nr. 1, S. 52–66.
- Boyer**, Kenneth K., John R. **Olson**, Roger J. **Calantone** und Eric C. **Jackson** (2002), Print Versus Electronic Surveys: A Comparison of Two Data Collection Methodologies, in: *Journal of Operations Management*, 20. Jg., Nr. 4, S. 357–373.
- Breckler**, Steven J. (1990), Applications of Covariance Structure Modeling in Psychology: Cause for Concern?, in: *Psychological Bulletin*, 52. Jg., S. 260–271.
- Brockner**, Joel (1992), The Escalation of Commitment to a Failing Course of Action: Toward Theoretical Progress, in: *Academy of Management Review*, 17. Jg., Nr. 1, S. 39–61.

- Brockner**, Joel und Jeffrey Z. **Rubin** (1985), *Entrapment in Escalating Conflicts*, Berlin: Springer.
- Brockner**, Joel, Jeffrey Z. **Rubin**, Judy **Fine**, Thomas P. **Hamilton**, Barbara **Thomas** und Beth **Turtsky** (1982), Factors Affecting Entrapment in Escalating Conflicts: The Importance of Timing, in: *Journal of Research in Personality*, 16. Jg., S. 247–266.
- Brockner**, Joel, Jeffrey Z. **Rubin** und Elaine **Lang** (1981), Face Saving and Entrapment, in: *Journal of Experimental Social Psychology*, 17. Jg., S. 68–79.
- Brockner**, Joel, Myril C. **Shaw** und Jeffrey Z. **Rubin** (1979), Factors Affecting Withdrawal from an Escalating Conflict: Quitting Before It's Too Late, in: *Journal of Experimental Social Psychology*, 15. Jg., S. 492–503.
- Brown**, Clifton E. und Ira **Solomon** (1993), An Experimental Investigation of Explanations for Outcome Effects on Appraisals of Capital-Budgeting Decisions, in: *Contemporary Accounting Research*, 10. Jg., Nr. 1, S. 83–111.
- Brown**, Roger L. (1994), Efficacy of the Indirect Approach for Estimating Structural Equation Models with Missing Data: A Comparison of Five Methods, in: *Structural Equation Modeling*, 1. Jg., S. 287–316.
- Browne**, Michael W. und Robert **Cudeck** (1989), Single Sample Cross-Validation Indices for Covariance Structures, in: *Multivariate Behavioral Research*, 24. Jg., S. 445–455.
- Browne**, Michael W. und Robert **Cudeck** (1993), Alternative Ways of Assessing Model Fit, in: K. A. Bollen und J. S. Long (Hrsg.), *Testing Structural Equation Models*, Newbury Park, CA: Sage, S. 136–162.
- Bruner**, Gordon C. und Paul J. **Hensel** (1992), *Marketing Scales Handbook: A Compilation of Multi-Item Measures*, Bd. 1, Chicago, IL: American Marketing Association.
- Bruner**, Gordon C. und Paul J. **Hensel** (1996), *Marketing Scales Handbook: A Compilation of Multi-Item Measures*, Bd. 2, Chicago, IL: American Marketing Association.
- Bruner II**, Gordon C., Paul J. **Hensel** und Karen E. **James** (2005), *Marketing Scales Handbook Volume IV: A Compilation of Multi-Item Measures for Consumer Behavior & Advertising*, Bd. 4, Chicago, IL: American Marketing Association.
- Bruner II**, Gordon C., Karen E. **James** und Paul J. **Hensel** (2001), *Marketing Scales Handbook: A Compilation of Multi-Item Scales*, Bd. 3, Chicago, IL: American Marketing Association.
- Bruner II**, Gordon C. und Maryon F. **King** (2003), Assessing Progress in Advertising-Related Scale Development & Usage, in: *Journal of Applied Business Research*, 19. Jg., Nr. 2, S. 97–107.
- Burnham**, Kenneth P. und David R. **Anderson** (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2. Aufl., New York, NY: Springer.
- Burns**, Tom und George M. **Stalker** (1961), *The Management of Innovation*, London: Tavistock.
- Buros Institute of Mental Measurements** (2007), *Mental Measurement Yearbook*, online verfügbar, URL: <http://buros.unl.edu/buros/jsp/search.jsp>; Download: 3. April 2007.

- Butler**, John K. (1991), Toward Understanding and Measuring Conditions of Trust: Evolution of a Conditions of Trust Inventory, in: *Journal of Management*, 17. Jg., S. 643–664.
- Byrne**, Barbara M. (2001), *Structural Equation Modeling With AMOS: Basic Concepts, Applications and Programming*, Mahwah, NJ: Erlbaum.
- Camerer**, Colin (1995), Individual Decision Making, in: J. H. Kagel und A. E. Roth (Hrsg.), *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press, Kapitel 8, S. 587–703.
- Camerer**, Colin F. (1999), The Econometrics and Behavioral Economics of Escalation of Commitment: A Re-examination of Staw and Hoang's NBA Data, in: *Journal of Economic Behavior & Organization*, 39. Jg., S. 59–82.
- Campbell**, Donald T. (1950), The Indirect Assessment of Social Attitudes, in: *Psychological Bulletin*, 47. Jg., Nr. 1, S. 15–38.
- Campbell**, Donald T. und Donald W. **Fiske** (1959), Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix, in: *Psychological Bulletin*, 56. Jg., S. 81–105.
- Campbell**, Tavis S., Mark D. **Holder** und Christopher R. **France** (2006), The Effects of Experimenter Status and Cardiovascular Reactivity on Pain Reports, in: *Pain*, 125. Jg., Nr. 3, S. 264–269.
- Caplan**, Edwin H. (1966), Behavioral Assumptions of Management Accounting, in: *The Accounting Review*, 41. Jg., Nr. 3, S. 496–509.
- Cheng**, Mandy M., Axel K.-D. **Schulz**, Peter F. **Luckett** und Peter **Booth** (2003), The Effects of Hurdle Rates on the Level of Escalation of Commitment in Capital Budgeting, in: *Behavioral Research in Accounting*, 15. Jg., S. 63–85.
- Churchill**, Gilbert A. Jr. (1979), A Paradigm for Developing Better Measures of Marketing Constructs, in: *Journal of Marketing Research*, 16. Jg., Februar, S. 64–73.
- Clemen**, Robert T. und Terence **Reilly** (2001), *Making Hard Decisions with DecisionTools[®]*, Pacific Grove, CA: Duxbury.
- Cliff**, Norman (1983), Some Cautions Concerning the Application of Causal Modeling Methods, in: *Multivariate Behavioral Research*, 18. Jg., S. 115–126.
- Cloyd**, C. Bryan und Brian C. **Spilker** (2000), Confirmation Bias in Tax Information Search: A Comparison of Law Students and Accounting Students, in: *Journal of the American Taxation Association*, 22. Jg., Nr. 2, S. 60–71.
- Coase**, Ronald H. (1984), The New Institutional Economics, in: *Journal of Institutional and Theoretical Economics*, 140. Jg., S. 299–231.
- Cohen**, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, NJ: Erlbaum.
- Cohen**, Jacob, Patricia **Cohen**, Stephen G. **West** und Leona S. **Aiken** (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3. Aufl., Mahwah, NJ: Erlbaum.

- Commes**, Max-Theodor und Richard **Lienert** (1983), Controlling im FuE-Bereich, in: Zeitschrift für Organisation, 52. Jg., Nr. 7, S. 347–354.
- Conlon**, Donald E. und Howard **Garland** (1993), The Role of Project Completion Information in Resource Allocation Decisions, in: Academy of Management Journal, 38. Jg., Nr. 2, S. 402–413.
- Conlon**, Edward J. und Marya L. **Leatherwood** (1989), Sunk Costs and Financial Decision Making: Integration and Implications, in: C. F. Lee (Hrsg.), Advances in Financial Planning and Forecasting, Bd. 3, Greenwich, CT: JAI Press, S. 37–61.
- Cook**, Colleen, Fred **Heath** und Russel L. **Thompson** (2000), A Meta-Analysis of Response Rates in Web- or Internet-based Surveys, in: Educational and Psychological Measurement, 60. Jg., S. 821–836.
- Cooper**, Arnold C., Carolyn Y. **Woo** und William C. **Dunkelberg** (1988), Entrepreneurs' Perceived Chances For Success, in: Journal of Business Venturing, 3. Jg., S. 97–108.
- Cooper**, Robert G., Scott J. **Edgett** und Elko J. **Kleinschmidt** (2002a), Optimizing the Stage-Gate Process: What Best-Practice Companies Do (Part I), in: Research in Technology Management, 45. Jg., S. 21–27.
- Cooper**, Robert G., Scott J. **Edgett** und Elko J. **Kleinschmidt** (2002b), Optimizing the Stage-Gate Process: What Best-Practice Companies Do (Part II), in: Research in Technology Management, 45. Jg., S. 43–50.
- Cosier**, Richard A. (1981), Dialectical Inquiry in Strategic Planning: A Case of Premature Acceptance?, in: Academy of Management Review, 6. Jg., Nr. 4, S. 643–648.
- Cote**, Joseph A. und M. Ronald **Buckley** (1987), Estimating Trait, Method, and Error Variance: Generalizing Across 70 Construct Validation Studies, in: Journal of Marketing Research, 24. Jg., Nr. 3, S. 315–318.
- Craig**, C. Samuel und John M. **McCann** (1978), Item Nonresponse in Mail Surveys: Extent and Correlates, in: Journal of Marketing Research, 15. Jg., Nr. 2, S. 285–289.
- Cronbach**, Lee J. (1951), Coefficient Alpha and the Internal Structure of Tests, in: Psychometrika, 16. Jg., Nr. 3, S. 297–317.
- Crowne**, Douglas P. und David **Marlowe** (1964), The Approval Motive: Studies in Evaluative Dependence, New York, NY: Wiley.
- Cyert**, Richard M. und James G. **March** (1963), Behavioral Theory of the Firm, Englewood Cliffs, NJ: Prentice Hall.
- d'Astous**, Alain und Lilia **Boujbel** (2007), Positioning Countries on Personality Dimensions: Scale Development and Implications for Country Marketing, in: Journal of Business Research, 60. Jg., S. 231–239.
- Davis**, Stan, F. Todd **DeZoort** und Lori S. **Kopp** (2006), The Effect of Obedience Pressure and Perceived Responsibility on Management Accountants' Creation of Budgetary Slack, in: Behavioral Research in Accounting, 18. Jg., S. 19–35.
- Dean**, Burton V. (1968), Evaluating, Selecting, and Controlling R&D Projects, Bd. 89 der Reihe AMA Research Study, o. O.: American Management Association, Inc.

- Dearborn**, Dewitt C. und Herbert A. **Simon (1958)**, Selective Perception: A Note on the Department Identification of Executives, in: *Sociometry*, 21. Jg., S. 140–144.
- Demleitner**, Klaus (2006), *Projekt-Controlling: Die kaufmännische Sicht der Projekte*, Bd. 14 der Reihe Westerham, Renningen: Expert-Verlag.
- Dempster**, Arthur P., Nan M. **Laird** und Donald B. **Rubin (1977)**, Maximum Likelihood Estimation From Incomplete Data via the EM Algorithm (with Discussion), in: *Journal of the Royal Statistical Society Series B*, 39. Jg., S. 1–38.
- Dent**, Arthur G. H. (1931), Budgetary Business Control in Practice, in: *The Manchester Guardian Commercial*, o. Jg., Nr. 24, S. 549.
- DeVellis**, Robert F. (2002), *Scale Development: Theory and Applications*, 2. Aufl., Bd. 26 der Reihe Applied Social Research Methods Series, Thousand Oaks, CA: Sage.
- Devine**, Carl T. (1960), Research Methodology and Accounting Theory Formation, in: *The Accounting Review*, 35. Jg., Nr. 3, S. 387–399.
- Devlin**, Susan J., H. K. **Dong** und Marbue **Brown (1993)**, Selecting a Scale for Measuring Quality, in: *Marketing Research*, 5. Jg., Nr. 3, S. 12–17.
- Diamantopoulos**, Adamantios (2005), The C-OAR-SE Procedure for Scale Development in Marketing: A Comment, in: *International Journal of Research in Marketing*, 22. Jg., S. 1–9.
- Diamantopoulos**, Adamantios und Heidi M. **Winklhofer (2001)**, Index Construction with Formative Indicators: An Alternative to Scale Development, in: *Journal of Marketing Research*, 38. Jg., S. 269–277.
- Die Kommission der Europäischen Gemeinschaften (2003)**, Empfehlung der Kommission vom 6. Mai 2003 betreffend die Definition der Kleinstunternehmen sowie der kleinen und mittleren Unternehmen, Amtsblatt der Europäischen Union, L 124, S. 36–41.
- Diller**, Hermann (2006), Probleme der Handhabung von Strukturgleichungsmodellen in der betriebswirtschaftlichen Forschung, in: *Die Betriebswirtschaft*, 6. Jg., S. 611–617.
- Dillman**, Don A. (1991), The Design and Administration of Mail Surveys, in: *Annual Review of Sociology*, 17. Jg., S. 225–249.
- Dillman**, Don A. (2000), *Mail and Internet Surveys: The Tailored Design Method*, 2. Aufl., New York, NY: Wiley.
- Dilts**, David M. und Ken R. **Pence (2006a)**, Executive Technical Survey, online verfügbar, URL: <http://www.rateyourrisk.org/execsurvey.htm>; Download: 13. Februar 2007.
- Dilts**, David M. und Ken R. **Pence (2006b)**, Impact of Role in the Decision to Fail: An Exploratory Study of Terminated Projects, in: *Journal of Operations Management*, 24. Jg., S. 378–396.
- Dilts**, David M. und Ken R. **Pence (2006c)**, Technical Project Manager Survey, online verfügbar, URL: <http://www.rateyourrisk.org/pmsurvey.htm>; Download: 13. Februar 2007.
- DIN 69901 (1987)**, *Projektwirtschaft; Projektmanagement; Begriffe*, Ausgabe 1987-08.

- Dionne**, Shelley D., Francis J. **Yammarino**, Leanne E. **Atwater** und Lawrence R. **James** (2002), Neutralizing Substitutes for Leadership Theory: Leadership Effects and Common-Source Bias, in: *Journal of Applied Psychology*, 87. Jg., Nr. 3, S. 454–464.
- Dreze**, Jacques H. (1958), *Individual Decision Making Under Partially Controllable Uncertainty*, New York, NY: Columbia University.
- Drummond**, Helga (1995), De-Escalation in Decision Making: A Case of a Disastrous Partnership, in: *Journal of Management Studies*, 32. Jg., Nr. 3, S. 265–281.
- Du**, Stephen, Mark **Keil**, Lars **Mathiassen**, Yide **Shen** und Amrit **Tiwana** (2007), Attention-shaping Tools, Expertise, and Perceived Control in IT Project Risk Assessment, in: *Decision Support Systems*, 43. Jg., S. 269–283.
- Dunn**, Steven C., Robert F. **Seaker** und Matthew A. **Waller** (1994), Latent Variables in Business Logistics Research: Scale Development and Validation, in: *Journal of Business Logistics*, 15. Jg., Nr. 2, S. 145–172.
- Dvir**, Dov, Tzvi **Raz** und Aaron J. **Shenhar** (2003), An Empirical Analysis of the Relationship Between Project Planning and Project Success, in: *International Journal of Project Management*, 21. Jg., Nr. 2, S. 89–95.
- Dyckman**, Thomas R. (1964), The Effects of Alternative Accounting Techniques on Certain Management Decisions, in: *Journal of Accounting Research*, 2. Jg., Nr. 1, S. 91–107.
- Educational Testing Service** (2007), ETS Test Link, online verfügbar, URL: <http://www.ets.org/testcoll/>; Download: 3. April 2007.
- Edwards**, Allen L. (1957), *The Social Desirability Variable in Personality Assessment and Research*, New York, NY: Holt, Rinehart, & Winston.
- Edwards**, Lynne K., Allen L. **Edwards** und Cheryl **Clark** (1988), Social Desirability and the Frequency of Social-Reinforcement Scale, in: *Journal of Personality and Social Psychology*, 54. Jg., Nr. 3, S. 526–529.
- Efron**, Bradley (1979), Bootstrap Methods: Another Look at the Jackknife, in: *Annals of Statistics*, 7. Jg., S. 1–26.
- Eggert**, Andreas und Georg **Fassott** (2005), Zur Verwendung formativer und reflektiver Indikatoren in Strukturgleichungsmodellen: Bestandsaufnahme und Anwendungsempfehlungen, in: *Handbuch PLS-Pfadmodellierung: Methode, Anwendung, Praxisbeispiele*, Stuttgart: Schäffer-Poeschel, S. 31–47.
- Ellingson**, Jill E., Paul R. **Sackett** und Leaetta M. **Hough** (1999), Social Desirability Corrections in Personality Measurement: Issues of Applicant Comparison and Construct Validity, in: *Journal of Applied Psychology*, 84. Jg., Nr. 2, S. 155–166.
- Ernst**, Holger (2001), Erfolgsfaktoren neuer Produkte: Grundlagen für eine valide empirische Forschung, Bd. 95 der Reihe Beiträge zur betriebswirtschaftlichen Forschung, Wiesbaden: DUV.
- Ernst**, Holger (2003a), Causes and Effects of a Single Informant Bias in Empirical Innovation Research, WHU-Forschungspapier Nr. 96, Vallendar: WHU – Otto Beisheim School of Management.

- Ernst, Holger (2003b)**, Ursachen eines Informant Bias und dessen Auswirkungen auf die Validität empirischer betriebswirtschaftlicher Forschung, in: Zeitschrift für Betriebswirtschaft, 73. Jg., Nr. 12, S. 1249–1275.
- Eurostat: Statistisches Amt der Europäischen Gemeinschaften (2007)**, Statistical Classification of Economic Activities in the European Community, Rev. 2, online verfügbar, URL: <http://ec.europa.eu/eurostat/ramon/>; Download: 16. April 2007.
- Ewing, Michael T. und Julie Napoli (2005)**, Developing and Validating a Multidimensional Nonprofit Brand Orientation Scale, in: Journal of Business Research, 58. Jg., S. 841–853.
- Fassott, Georg (2006)**, Operationalisierung latenter Variablen in Strukturgleichungsmodellen, in: Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 58. Jg., S. 67–88.
- Favre, Anne-Catherine, Alina Matei und Yves Tillé (2005)**, Calibrated Random Imputation For Qualitative Data, in: Journal of Statistical Planning and Inference, 128. Jg., S. 411–425.
- Ferguson, William B. (1920)**, Relationship of Cost Accounting to Business Management: From the Viewpoint of the Business Executive, in: NACA Yearbook, o. O.: National Association of Cost Accountants.
- Festinger, Leon (1957)**, Theory of Cognitive Dissonance, Evanston, IL: Row, Peterson.
- Festinger, Leon und James M. Carlsmith (1959)**, Cognitive Consequences of Forced Compliance, in: Journal of Abnormal and Social Psychology, 58. Jg., S. 203–211.
- Festinger, Leon, Henry Riecken und Stanley Schachter (1956)**, When Prophecy Fails, New York, NY: Harper & Row.
- Fichman, Robert G., Mark Keil und Amrit Tiwana (2005)**, Beyond Valuation: Options Thinking in IT Project Management, in: California Management Review, 47. Jg., Nr. 2, S. 74–96.
- Fischhoff, Baruch (1975)**, Hindsight \neq Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty, in: Journal of Experimental Psychology: Human Perception and Performance, o. Jg., S. 288–299.
- Fisher, Robert J. (1993)**, Social Desirability Bias and the Validity of Indirect Questioning, in: Journal of Consumer Research, 20. Jg., Nr. 2, S. 303–315.
- Fisher, Robert J. (2000)**, The Future of Social-Desirability Bias Research in Marketing, in: Psychology & Marketing, 17. Jg., Nr. 2, S. 73–77.
- Fisher, Robert J. und James E. Katz (2000)**, Social-Desirability Bias and the Validity of Self-Reported Values, in: Psychology & Marketing, 17. Jg., Nr. 2, S. 105–120.
- Flynn, Leisa R. und Dawn Percy (2001)**, Four Subtle Sins in Scale Development: Some Suggestions for Strengthening the Current Paradigm, in: International Journal of Market Research, 43. Jg., Nr. 4, S. 409–423.
- Fornell, Claes und David F. Larcker (1981)**, Evaluating Structural Equations Models with Unobservable Variables and Measurement Error, in: Journal of Marketing Research, 18. Jg., S. 39–50.

- Fowler**, Floyd J., Anthony M. **Roman** und Zhu X. **Di** (1998), Mode Effect in a Survey of Medicare Prostate Surgery Patients, in: *Public Opinion Quarterly*, 62. Jg., S. 29–46.
- Fox**, Shaul, Amalia **Schmida** und Yoel **Yinon** (1995), Escalation Behavior in Domains Related and Unrelated to Decision Makers' Academic Background, in: *Journal of Business and Psychology*, 10. Jg., Nr. 2, S. 245–259.
- Fraley**, R. Chris (2004), *How to Conduct Behavioral Research Over the Internet: A Beginner's Guide to HTML and CGI/Perl*, New York, NY: Guilford Press.
- Friedman**, Milton (1971), The Methodology of Positive Economics, in: W. Breit und H. Hochman (Hrsg.), *Readings in Microeconomics*, 2. Aufl., Hinsdale, IL: Dryden Press, S. 23–54.
- Galbraith**, Jay R. (1973), *Designing Complex Organizations*, Reading, MA: Addison-Wesley.
- Garland**, Howard, Craig A. **Sandefur** und Anne C. **Rogers** (1990), De-Escalation of Commitment in Oil Exploration: When Sunk Costs and Negative Feedback Coincide, in: *Journal of Applied Psychology*, 75. Jg., Nr. 6, S. 721–727.
- Garver**, Michael S. und John T. **Mentzer** (1999), Logistics Research Methods: Employing Structural Equation Modeling to Test for Construct Validity, in: *Journal of Business Logistics*, 20. Jg., S. 33–57.
- Gerbing**, David W. und James C. **Anderson** (1988), An Updated Paradigm for Scale Development Incorporating Unidimensionality and its Assessment, in: *Journal of Marketing Research*, 25. Jg., Nr. 2, S. 186–192.
- Gerling**, Patrick G. (2007), *Controlling und Kognition*, Bd. 6 der Reihe *Controlling*, Lohmar: Josef Eul Verlag.
- Gervais**, Simon und Terrance **Odean** (2001), Learning to Be Overconfident, in: *The Review of Financial Studies*, 14. Jg., Nr. 1, S. 1–27.
- Ghosh**, Dipankar (1995), Throwing Good Money After Bad, in: *Management Accounting*, o. Jg., Juli, S. 51–54.
- Ghosh**, Dipankar (1997), De-escalation Strategies: Some Experimental Evidence, in: *Behavioral Research in Accounting*, 9. Jg., S. 88–112.
- Giere**, Jens, Bernd W. **Wirtz** und Oliver **Schilke** (2006), Mehrdimensionale Konstrukte: Konzeptionelle Grundlagen und Möglichkeiten ihrer Analyse mithilfe von Strukturgleichungsmodellen, in: *Die Betriebswirtschaft*, 66. Jg., Nr. 6, S. 678–695.
- Gilley**, Otis W. und Robert P. **Leone** (1991), A Two-Stage Imputation Procedure for Item Non-response in Surveys, in: *Journal of Business Research*, 22. Jg., S. 281–291.
- Goethals**, George R. (1992), Dissonance and Self-Justification, in: *Psychological Inquiry*, 3. Jg., Nr. 4, S. 327–329.
- Goethals**, George R. und Richard F. **Reckman** (1973), The Perception of Consistency in Attitudes, in: *Journal of Experimental Social Psychology*, 9. Jg., Nr. 6, S. 491–501.
- Graham**, John W., Scott M. **Hofer** und David P. **MacKinnon** (1996), Maximizing the Usefulness of Data Obtained with Planned Missing Value Patterns: An Application of Maximum Likelihood Procedures, in: *Multivariate Behavioral Research*, 31. Jg., Nr. 2, S. 197–218.

- Grewal**, Rajdeep, Joseph **Cote** und Hans **Baumgartner** (2004), Multicollinearity and Measurement in Structural Equation Models: Implications for Theory Testing, in: *Marketing Science*, 23. Jg., S. 519–529.
- Grieshop**, Hendrik (2008), Kooperation von Controllerebereich und externem Rechnungswesen: Messung, Wirkung, Determinanten, Vallendar: Dissertation in Vorbereitung.
- Griffis**, Stanley E., Thomas J. **Goldsby** und Martha **Cooper** (2003), Web-Based and Mail Surveys: A Comparison of Response, Data, and Cost, in: *Journal of Business Logistics*, 24. Jg., Nr. 2, S. 237–258.
- Griffith**, David A. und Michael G. **Harvey** (2001), A Resource Perspective of Global Dynamic Capabilities, in: *Journal of International Business Studies*, 32. Jg., Nr. 3, S. 597–606.
- Guadagnoli**, Edward und Wayne F. **Velicer** (1988), Relation of Sample Size to the Stability of Component Patterns, in: *Psychological Bulletin*, 103. Jg., S. 461–468.
- Guilford**, Joy P. (1954), *Psychometric Methods*, 2. Aufl., New York, NY: McGraw-Hill.
- Guler**, Isin (2007), Throwing Good Money After Bad? Political and Institutional Influences on Sequential Decision Making in the Venture Capital Industry, in: *Administrative Science Quarterly*, 52. Jg., S. 248–285.
- Hague**, Paul (1993), *Questionnaire Design*, London: Kogan Page Limited.
- Hair**, Joseph, Robert **Bush** und Dave **Ortinou** (2000), *Marketing Research: A Practical Approach for the New Millennium*, Columbus: McGraw-Hill.
- Hall**, Peter (1982), *Great Planning Disasters*, Berkeley, CA: University of California Press.
- Hammond**, John S., Ralph L. **Keeney** und Howard **Raiffa** (2006), The Hidden Traps in Decision Making, in: *Harvard Business Review*, 84. Jg., Nr. 1, S. 118–126.
- Hands**, D. Wade (2001), *Reflection without Rules: Economic Methodology and Contemporary Science Theory*, Cambridge, UK: Cambridge University Press.
- Hantula**, Donald A. und Jennifer L. DeNicolis **Bragger** (1999), The Effects of Feedback Equivocality on Escalation of Commitment: An Empirical Investigation of Decision Dilemma Theory, in: *Journal of Applied Social Psychology*, 29. Jg., Nr. 2, S. 424–444.
- Hardesty**, David M. und William O. **Bearden** (2004), The Use of Expert Judges in Scale Development: Implications for Improving Face Validity of Measures of Unobservable Constructs, in: *Journal of Business Research*, 57. Jg., S. 98–107.
- Harris**, Peter und Wendy **Middleton** (1994), The Illusion of Control and Optimism about Health: On Being Less at Risk but no more in Control Than Others, in: *British Journal of Social Psychology*, 33. Jg., S. 369–386.
- Harrison**, Paul D., Chee W. **Chow**, Anne **Wu** und Adrian M. **Harrell** (1999), A Cross-Cultural Investigation of Managers' Project Evaluation Decisions, in: *Behavioral Research in Accounting*, 11. Jg., S. 143–160.
- Hartig**, Johannes, Andreas **Frey** und Nina **Jude** (2007), Validität, in: H. Moosbrugger und A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*, Heidelberg: Springer, Kapitel 7, S. 136–163.

- Hasher**, Lynn, M. S. **Attig** und J. W. **Alba** (1981), I Knew It All Along: Or, Did I?, in: *Journal of Verbal Learning and Verbal Behaviour*, 20. Jg., S. 86–96.
- Hauschildt**, Jürgen (2004), *Innovationsmanagement*, 3. Aufl., München: Vahlen.
- Hawkins**, Scott A. und Reid **Hastie** (1990), Hindsight: Biased Judgments of Past Events after the Outcomes Are Known, in: *Psychological Bulletin*, 107. Jg., S. 311–327.
- Heath**, Chip (1995), Escalation and De-Escalation of Commitment in Response to Sunk Costs: The Role of Budgeting in Mental Accounting, in: *Organizational Behavior and Human Decision Processes*, 62. Jg., Nr. 1, April, S. 38–54.
- Heine**, Bernd-O., Bernhard **Hirsch**, Klaus **Hufschlag**, Marc **Lesch**, Matthias **Meyer**, Roman **Müller**, Anne **Paefgen** und Guido **Pieroth** (2006), Zur Modellierung ökonomischer Akteure mit begrenzten kognitiven Fähigkeiten: Anleitung zu einer problemspezifischen Ausdifferenzierung des Homo oeconomicus, WHU-Forschungspapier Nr. 110, Vallendar: WHU – Otto Beisheim School of Management.
- Helasoja**, Ville, Ritva **Prättälä**, Liudmila **Dregval**, Iveta **Pudule** und Anu **Kasmel** (2002), Late Response and item Nonresponse in the Finbalt Health Monitor Survey, in: *European Journal of Public Health*, 12. Jg., S. 117–123.
- Heng**, Cheng-Suang, Bernard C. Y. **Tan** und Kwok-Kee **Wei** (2003), De-escalation of Commitment in Software Projects: Who Matters? What Matters?, in: *Information & Management*, 41. Jg., S. 99–110.
- Hensley**, Rhonda L. (1999), A Review of Operations Management Studies Using Scale Development Techniques, in: *Journal of Operations Management*, 17. Jg., S. 343–358.
- Hildebrandt**, Lutz und Dirk **Temme** (2006), Probleme der Validierung mit Strukturgleichungsmodellen, in: *Die Betriebswirtschaft*, 66. Jg., S. 618–639.
- Hinkin**, Timothy R. (1995), A Review of Scale Development Practices in the Study of Organizations, in: *Journal of Management*, 21. Jg., Nr. 5, S. 967–988.
- Hippel**, Paul T. von (2004), Biases in SPSS 12.0 Missing Value Analysis, in: *The American Statistician*, 58. Jg., Nr. 2, S. 160–164.
- Hirsch**, Bernhard (2007), *Controlling und Entscheidungen: Zur verhaltenswissenschaftlichen Fundierung des Controlling*, Tübingen: Mohr Siebeck.
- Hoelter**, Jon W. (1983), The Analysis of Covariance Structures: Goodness-of-Fit Indices, in: *Sociological Methods and Research*, 11. Jg., S. 325–344.
- Hofstedt**, Thomas R. und James C. **Kinard** (1970), A Strategy for Behavioral Accounting Research, in: *The Accounting Review*, 45. Jg., Nr. 1, S. 38–54.
- Hollar**, David W. (1997), *Escalation Behavior in Higher Education: A Structural Equation Model*, Greensboro: The University of North Carolina at Greensboro.
- Hollar**, David W., John **Hattie**, Bert **Goldman** und James **Lancaster** (2000), Developing Assessment Procedures and Assessing Two Models of Escalation Behavior Among Community College Administrators, in: *Theory and Decision*, 49. Jg., Nr. 1, August, S. 1–24.

- Holmes, David S. (1968)**, Dimensions of Projection, in: *Psychological Bulletin*, 69. Jg., April, S. 248–268.
- Homann, Karl und Andreas Suchanek (2000)**, *Ökonomik: Eine Einführung*, Tübingen: Mohr Siebeck.
- Homburg, Christian (1992a)**, Beispiele zur Kausalanalyse, in: *Wirtschaftswissenschaftliches Studium*, 10. Jg., S. 541–544.
- Homburg, Christian (1992b)**, Die Kausalanalyse: Eine Einführung, in: *Wirtschaftswissenschaftliches Studium*, 21. Jg., Nr. 10, S. 499–508.
- Homburg, Christian und Hans Baumgartner (1995)**, Beurteilung von Kausalmodellen: Bestandsaufnahme und Anwendungsempfehlungen, in: *Marketing – Zeitschrift für Forschung und Praxis*, 17. Jg., Nr. 3, S. 162–176.
- Homburg, Christian und Hans Baumgartner (1998)**, Beurteilung von Kausalmodellen: Bestandsaufnahme und Anwendungsempfehlungen, in: L. Hildebrandt und C. Homburg (Hrsg.), *Die Kausalanalyse: Ein Instrument der empirischen betriebswirtschaftlichen Forschung*, Stuttgart: Schäffer-Poeschel, S. 343–369.
- Homburg, Christian und Andreas Dobratz (1991)**, Iterative Modellselektion in der Kausalanalyse, in: *Zeitschrift für betriebswirtschaftliche Forschung*, 43. Jg., Nr. 3, S. 213–237.
- Homburg, Christian und Annette Giering (1996)**, Konzeptionalisierung und Operationalisierung komplexer Konstrukte, in: *Marketing – Zeitschrift für Forschung und Praxis*, 18. Jg., Nr. 1, S. 5–24.
- Homburg, Christian und Martin Klarmann (2006)**, Die Kausalanalyse in der empirischen betriebswirtschaftlichen Forschung: Problemfelder und Anwendungsempfehlungen, in: *Die Betriebswirtschaft*, 66. Jg., S. 727–748.
- Homburg, Christian und Christian Pflessner (2000)**, Strukturgleichungsmodelle mit latenten Variablen: Kausalanalyse, in: A. Herrmann und C. Homburg (Hrsg.), *Marktforschung: Methoden, Anwendungen, Praxisbeispiele*, 2. Aufl., Wiesbaden: Gabler.
- Homburg, Christian und Ruth Stock (2004)**, The Link Between Salespeople’s Job Satisfaction and Customer Satisfaction in a Business-to-Business Context: A Dyadic Analysis, in: *Journal of the Academy of Marketing Science*, 32. Jg., S. 144–158.
- Hoppenstedt Firmendatenbank (2007)**, 2.1 (Relaunch Update), Stand: Juni 2007.
- Hu, Li-tze und Peter M. Bentler (1995)**, Evaluating Model Fit, in: R. H. Hoyle (Hrsg.), *Structural Equation Modeling: Concepts, Issues, and Applications*, Thousand Oaks, CA: Sage, Kapitel 5, S. 76–99.
- Hu, Li-tze und Peter M. Bentler (1998)**, Fit Indices in Covariance Structure Analysis: Sensitivity to Underparameterized Model Misspecification, in: *Psychological Methods*, 3. Jg., S. 424–453.
- Hu, Li-tze und Peter M. Bentler (1999)**, Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives, in: *Structural Equation Modeling*, 6. Jg., S. 1–55.

- Hubner**, Elmar (1996), Abbruchentscheidungen im F&E-Management, Wien: Wirtschaftsuniversität Wien.
- Hughes**, Shawn W., Donald D. **Tippett** und Warren K. **Thomas** (2004), Measuring Project Success in the Construction Industry, in: Engineering Management Journal, 16. Jg., Nr. 3, S. 31–37.
- Hummel**, Siegfried und Wolfgang **Männel** (1986), Kostenrechnung 1: Grundlagen, Aufbau und Anwendung, Wiesbaden: Gabler.
- Hunter**, John E. (2001), The Desperate Need for Replications, in: Journal of Consumer Research, 28. Jg., Nr. 1, S. 149–158.
- Ichikawa**, Masanori und Sadanori **Konishi** (1995), Application of the Bootstrap Methods in Factor Analysis, in: Psychometrika, 60. Jg., S. 77–93.
- Jackson**, Paul R., Toby D. **Wall**, Robin **Martin** und Keith **Dauids** (1993), New Measures of Job Control, Cognitive Demand, and Production Responsibility, in: Journal of Applied Psychology, 78. Jg., S. 753–762.
- Jacoby**, Jacob (1978), Consumer Research: How Valid and Useful are all our Consumer Behavior Research Findings? A State of the Art Review, in: Journal of Marketing, 42. Jg., S. 87–96.
- Jamal**, Karim (1993), Comments on “An Experimental Investigation of Explanations for Outcome Effects on Appraisals of Capital-Budgeting Decisions”, in: Contemporary Accounting Research, 10. Jg., Nr. 1, S. 113–117.
- Janis**, Irving L. und Leon **Mann** (1977), Decision Making: A Psychological Analysis of Conflict, Choice and Commitment, New York, NY: Free Press.
- Jarvis**, Cheryl B., Scott B. **MacKenzie** und Philip M. **Podsakoff** (2003), A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research, in: Journal of Consumer Research, 30. Jg., September, S. 199–218.
- Jo**, Myung-Soo (2000), Controlling Social-Desirability Bias via Method Factors of Direct and Indirect Questioning in Structural Equation Models, in: Psychology & Marketing, 17. Jg., Nr. 2, S. 137–148.
- Johnson**, Jim (1995), Chaos: The Dollar Drain of IT Project Failures, in: Application Development Trends, 2. Jg., Nr. 1, S. 41–47.
- Johnson**, Jim (2000), Turning Chaos into Success, in: Software Magazine, 19. Jg., S. 30–39.
- Jones**, Malcolm M. und Ephraim R. **McLean** (1970), Management Problems in Large-Scale Software Development Projects, in: Industrial Management Review, 11. Jg., Nr. 3, S. 1–15.
- Jost**, Peter-J. (2000), Organisation und Motivation: Eine ökonomisch-psychologische Einführung, Wiesbaden: Gabler.
- Jöreskog**, Karl G. und Dag **Sörbom** (1981), LISREL V: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Squares Methods (Research Report 81-8), Uppsala, Schweden: University of Uppsala, Department of Statistics.

- Jöreskog**, Karl G. und Dag **Sörbom** (1989), LISREL 7 Users's Reference Guide, Chicago, IL: SPSS Publications.
- Jungermann**, Helmut, Hans-Rüdiger **Pfister** und Katrin **Fischer** (1998), Die Psychologie der Entscheidung: Eine Einführung, Heidelberg: Spektrum Akademischer Verlag.
- Kadous**, Kathryn und Lisa M. **Sedor** (2004), The Efficacy of Third-Party Consultation in Preventing Managerial Escalation of Commitment: The Role of Mental Representations, in: Contemporary Accounting Research, 21. Jg., Nr. 1, S. 55–82.
- Kahn**, Kenneth B. (1996), Interdepartmental Integration: A Definition with Implications for Product Development Performance, in: Journal of Product Innovation Management, 13. Jg., Nr. 1, S. 137–151.
- Kalton**, Graham (1983), Introduction to Survey Sampling, Bd. 07-035 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Newbury Park, CA: Sage Publications.
- Kaplan**, David (1990), Evaluating and Modifying Covariance Structure Models: A Review and Recommendation, in: Multivariate Behavioral Research, 25. Jg., S. 137–155.
- Keil**, Mark (1995), Pulling the Plug: Software Project Management and the Problem of Project Escalation, in: Management Information Systems Quarterly, 19. Jg., Nr. 4, S. 421–447.
- Keil**, Mark, Richard **Mixon**, Timo **Saarinen** und Virpi **Tuunainen** (1995), Understanding Run-away Information Technology Projects: Results from an International Research Programm Based on Escalation Theory, in: Journal of Management Information Systems, 11. Jg., Nr. 3, S. 65–85.
- Keil**, Mark, Arun **Rai** und Joan **Mann** (2000), Why Software Projects Escalate: An Empirical Analysis and Test of Four Theoretical Models, in: Management Information Systems Quarterly, 24. Jg., Nr. 4, S. 631–664.
- Keil**, Mark, Arun **Rai**, Joan E. C. **Mann** und G. Peter **Zhang** (2003), Why Software Projects Escalate: The Importance of Project Management Constructs, in: IEEE Transactions on Engineering Management, 50. Jg., Nr. 3, August, S. 251–261.
- Keil**, Mark und Daniel **Robey** (1999), Turning Around Troubled Software Projects: An Exploratory Study of the Deescalation of Commitment to Failing Courses of Action, in: Journal of Management Information Systems, 15. Jg., Nr. 4, S. 63–87.
- Keil**, Mark und Daniel **Robey** (2001), Blowing the Whistle on Troubled Software Projects, in: Communications of the Association for Computing Machinery, 44. Jg., Nr. 4, S. 87–93.
- Keil**, Mark, Bernard C. Y. **Tan**, Kwok-Kee **Wei**, Timo **Saarinen**, Virpi **Tuunainen** und Arjen **Wassenaar** (2000), A Cross-Cultural Study on Escalation of Commitment Behavior in Software Projects, in: Management Information Systems Quarterly, 24. Jg., Nr. 2, S. 299–325.
- Keil**, Mark, Duane P. **Truex III** und Richard **Mixon** (1995), The Effects of Sunk Costs and Project Completion on Information Technology Project Escalation, in: IEEE Transactions on Engineering Management, 42. Jg., Nr. 4, S. 372–381.
- Keller**, Scott B., Katrina **Savitskie**, Theodore P. **Stank**, Daniel F. **Lynch** und Alexander E. **Ellinger** (2002), A Summary and Analysis of Multi-Item Scales Used in Logistics Research, in: Journal of Business Logistics, 23. Jg., Nr. 2, S. 83–281.

- Kellogg**, Ronald T. (2002), *Cognitive Psychology*, 2. Aufl., Thousand Oaks, CA: Sage.
- Kernan**, Mary C. und Robert G. **Lord** (1989), The Effects of Explicit Goals and Specific Feedback on Escalation Processes, in: *Journal of Applied Social Psychology*, 19. Jg., Nr. 13, S. 1125–1143.
- Kiesler**, Charles A. und Joseph **Sakumura** (1966), A Test of a Model for Commitment, in: *Journal of Personality and Social Psychology*, 3. Jg., Nr. 3, S. 349–353.
- Kim**, Chankon und Hanjoon **Lee** (1997), Development of Family Triadic Measures for Children's Purchase Influence, in: *Journal of Marketing Research*, 34. Jg., Nr. 3, S. 307–321.
- King**, Maryon F. und Gordon C. **Bruner** (2000), Social Desirability Bias: A Neglected Aspect of Validity Testing, in: *Psychology & Marketing*, 17. Jg., Nr. 2, S. 79–103.
- Kirby**, Susan L. und Mark A. **Davis** (1998), A Study of Escalating Commitment in Principal-Agent Relationships: Effects of Monitoring and Personal Responsibility, in: *Journal of Applied Psychology*, 83. Jg., Nr. 2, S. 206–217.
- Kisfalvi**, Veronika (2000), The Threat of Failure, the Perils of Success and CEO Character: Sources of Strategic Persistence, in: *Organization Studies*, 21. Jg., Nr. 3, S. 611–639.
- Kline**, Rex B. (2005), *Principles and Practice of Structural Equation Modeling*, 2. Aufl., New York, NY: Guilford Press.
- Knight**, Frank H. (1971), *Risk, Uncertainty and Profit*, Nachdruck der Erstauflage von 1921, Chicago, IL: University of Chicago Press.
- Koertge**, Noretta (1975), Popper's Metaphysical Research Programme for the Human Sciences, in: *Inquiry*, 19. Jg., S. 437–462.
- Koertge**, Noretta (1979), The Methodological Status of Popper's Rationality Principle, in: *Theory and Decision*, 10. Jg., S. 83–95.
- Korman**, Abraham K. (1974), Contingency Approaches to Leadership: An Overview, in: J. G. Hunt und L. L. Larson (Hrsg.), *Contingency Approaches to Leadership*, Carbondale: Southern Illinois University Press, S. 189–195.
- Kurzban**, Robert, Peter **DeScioli** und Erin **O'Brien** (2007), Audience Effects on Moralistic Punishment, in: *Evolution and Human Behavior*, 28. Jg., Nr. 2, S. 75–84.
- Kwok**, Winston C. C. und David J. **Sharp** (1998), A Review of Construct Measurement Issues in Behavioral Accounting Research, in: *Journal of Accounting Literature*, 17. Jg., S. 137–174.
- Kwon**, Hyungil und Galen **Trail** (2005), The Feasibility of Single-Item Measures in Sport Loyalty Research, in: *Sport Management Review*, 8. Jg., S. 69–89.
- Lange**, Edgar C. (1993), *Abbruchentscheidungen bei F&E-Projekten*, Wiesbaden: DUV.
- Lant**, Theresa K. und Amy E. **Hurley** (1999), A Contingency Model of Response to Performance Feedback: Escalation of Commitment and Incremental Adaptation in Resource Investment Decisions, in: *Group & Organization Management*, 24. Jg., Nr. 4, S. 421–437.
- Laux**, Helmut (2005), *Entscheidungstheorie*, 6. Aufl., Berlin: Springer.

- Lawrence**, Paul und Jay **Lorsch** (1967), *Organization and Environment: Managing Differentiation and Integration*, Boston, MA: Harvard Business School Press.
- Lee**, Eun S. und Ronald N. **Forthofer** (2006), *Analyzing Complex Survey Data*, 2. Aufl., Bd. 07-071 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.
- Lichtenstein**, Sarah und Baruch **Fischhoff** (1977), Do Those Who Know More Also Know More About How Much They Know?: The Calibration of Probability Judgments, in: *Organizational Behavior and Human Performance*, 20. Jg., Nr. 2, S. 159–183.
- Lichtenstein**, Sarah, Baruch **Fischhoff** und Lawrence D. **Phillips** (1982), Calibration of Probabilities: The State of the Art to 1980, in: D. Kahneman, P. Slovic und A. Tversky (Hrsg.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge, UK: Cambridge University Press, S. 306–334.
- Likert**, Rensis (1969), *The Human Organization: Its Management and Value*, New York, NY: McGraw-Hill.
- Lindell**, Michael K. und David J. **Whitney** (2001), Accounting for Common Method Variance in Cross-Sectional Research Design, in: *Journal of Applied Psychology*, 86. Jg., Nr. 1, S. 114–121.
- Lindenberg**, Siegwart (1985), An Assessment of the New Political Economy: Its Potential for the Social Sciences and for Sociology in Particular, in: *Sociological Theory*, 3. Jg., Nr. 1, S. 99–114.
- Lindenberg**, Siegwart (1990), Homo Socio-oeconomicus: The Emergence of a General Model of Man in the Social Sciences, in: *Journal of Institutional and Theoretical Economics*, 146. Jg., S. 727–748.
- Lindenberg**, Siegwart (1991), Die Methode der abnehmenden Abstraktion: Theoriegesteuerte Analyse und empirischer Gehalt, in: H. Esser und K. G. Troitzsch (Hrsg.), *Modellierung sozialer Prozesse*, Bd. 2 der Reihe Sozialwissenschaftliche Tagungsberichte, Bonn: Informationszentrum Sozialwissenschaften, S. 29–73.
- Lindhe**, Richard (1963), Accelerated Depreciation for Income Tax Purposes – A Study of the Decision and Some Firms Who Made It, in: *Journal of Accounting Research*, 1. Jg., Nr. 2, S. 139–148.
- Lingnau**, Volker (2006), *Controlling - Ein kognitionsorientierter Ansatz*, 2. Aufl., Bd. 4 der Reihe Beiträge zur Controlling-Forschung, Kaiserslautern: Lehrstuhl für Unternehmensrechnung und Controlling, Technische Universität Kaiserslautern.
- Lissitz**, Robert W. und Samuel B. **Green** (1975), Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach, in: *Journal of Applied Psychology*, 60. Jg., S. 10–13.
- Little**, Roderick J. A. (1988), A Test of Missing Completely at Random for Multivariate Data With Missing Values, in: *Journal of the American Statistical Association*, 83. Jg., Nr. 404, S. 1198–1202.
- Little**, Roderick J. A. und Donald B. **Rubin** (1989), The Analysis of Social Science Data With Missing Values, in: *Sociological Methods and Research*, 18. Jg., S. 292–326.

- Livingstone**, John L. (1967), A Behavioral Study of Tax Allocation in Electric Utility Regulation, in: *The Accounting Review*, 42. Jg., Nr. 3., S. 544–552.
- Lord**, Alan T. (1989), The Development of Behavioral Thought in Accounting, 1952-1981, in: *Behavioral Research in Accounting*, 1. Jg., S. 124–149.
- Lovallo**, Dan und Daniel **Kahneman** (2003), Delusions of Success: How Optimism Undermines Executives' Decisions, in: *Harvard Business Review*, 8. Jg., Nr. 1, S. 56–63.
- Lueg**, Jason E. und R. Zachary **Finney** (2007), Interpersonal Communication in the Consumer Socialization Process: Scale Development and Validation, in: *Journal of Marketing Theory and Practice*, 15. Jg., Nr. 1, S. 25–39.
- Luft**, Joan und Michael D. **Shields** (2003), Mapping Management Accounting: Graphics and Guidelines for Theory-Consistent Empirical Research, in: *Accounting, Organizations and Society*, 28. Jg., S. 169–249.
- MacKenzie**, Scott B., Philip M. **Podsakoff** und Richard **Fetter** (1991), Organizational Citizenship Behavior and Objective Productivity as Determinants of Managerial Evaluations of Salespersons' Performance, in: *Organizational Behavior and Human Decision Processes*, 50. Jg., S. 123–150.
- Mahoney**, Joseph T. und J. Rajendran **Pandian** (1992), The Resource-Based View Within the Conversation of Strategic Management, in: *Strategic Management Journal*, 13. Jg., S. 363–380.
- Maines**, Lauren A., Gerald L. **Salamon** und Geoffrey B. **Sprinkle** (2006), An Information Economic Perspective on Experimental Research in Accounting, in: *Behavioral Research in Accounting*, 18. Jg., S. 85–102.
- Malhotra**, Naresh K., Sung S. **Kim** und Ashutosh **Patil** (2006), Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research, in: *Management Science*, 52. Jg., Nr. 12, S. 1865–1883.
- Mansfield**, Edwin, John **Rapoport**, Jerome **Schnee**, Sam **Wagner** und Michael **Hamburger** (1971), *Research and Innovation in the Modern Corporation*, New York, NY: W. W. Norton.
- March**, James G. und Zur **Shapira** (1987), Managerial Perspectives on Risk and Risk Taking, in: *Management Science*, 33. Jg., S. 1404–1418.
- March**, James G. und Herbert A. **Simon** (1958), *Organizations*, New York, NY: Wiley.
- Marsh**, Herbert W., John R. **Balla** und Roderick P. **McDonald** (1988), Goodness-of-Fit in Confirmatory Factor Analysis: The Effect of Sample Size, in: *Psychological Bulletin*, 103. Jg., Nr. 3, S. 391–410.
- McCain**, Bruce E. (1986), Continuing Investment Under Conditions of Failure: A Laboratory Study of the Limits to Escalation, in: *Journal of Applied Psychology*, 71. Jg., Nr. 2, S. 280–284.
- McFadden**, Daniel (1999), Rationality for Economists?, in: *Journal of Risk and Uncertainty*, 19. Jg., Nr. 1-3, S. 73–105.

- McGrath**, Rita G., Walter J. **Ferrier** und Aubrey L. **Mendelow** (2004), Real Options as Engines of Choice and Heterogeneity, in: *Academy of Management Review*, 29. Jg., Nr. 1, S. 86–101.
- Meckling**, William H. (1976), Values and the Choice of the Model of the Individual in the Social Sciences, in: *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 112. Jg., S. 545–559.
- Meglino**, Bruce M. und M. Audrey **Korsgaard** (2004), Considering Rational Self-Interest as a Disposition: Organizational Implications of Other Orientation, in: *Journal of Applied Psychology*, 89. Jg., Nr. 6, S. 946–959.
- Mensch**, Barbara S. und Denise B. **Kandel** (1988), Underreporting of Substance Use in a National Longitudinal Youth Cohort, in: *Public Opinion Quarterly*, 52. Jg., S. 100–124.
- Meredith**, Jack (1988), Project Monitoring for Early Termination, in: *Project Management Journal*, 19. Jg., Nr. 5, S. 31–38.
- Messick**, Samuel (1989), Validity, in: R. L. Linn (Hrsg.), *Educational Measurement*, 3. Aufl., New York, NY: American Council on Education and National Council on Measurement in Education, S. 13–103.
- Meyer**, Matthias (2008), Integration von Kognition und Anreizen in die Rechnungswesenforschung, Habilitationsschrift, Vallendar: WHU – Otto Beisheim School of Management.
- Meyer**, Michael und John T. **Rigsby** (2001), A Descriptive Analysis of the Content and Contributors of Behavioral Research in Accounting 1989-1998, in: *Behavioral Research in Accounting*, 13. Jg., S. 253–278.
- Mick**, David G. (1996), Are Studies of Dark Side Variables Confounded by Socially Desirable Responding? The Case of Materialism, in: *Journal of Consumer Research*, 23. Jg., S. 329–353.
- Middleton**, Karen L. und Jeri L. **Jones** (2000), Social Desirable Response Sets: The Impact of Country Culture, in: *Psychology & Marketing*, 17. Jg., Nr. 2, S. 149–163.
- Min**, Soonhong, John T. **Mentzer** und Robert T. **Ladd** (2007), A Market Orientation in Supply Chain Management, in: *Journal of the Academy of Marketing Science*, 35. Jg., S. 507–522.
- Mohr**, Lawrence B. (1990), Understanding Significance Testing, Bd. 07-073 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Newbury Park, CA: Sage.
- Montealegre**, Ramiro und Mark **Keil** (2000), De-Escalating Information Technology Projects: Lessons from the Denver International Airport, in: *Management Information Systems Quarterly*, 24. Jg., Nr. 3, September, S. 417–447.
- Moon**, Henry (2001a), Looking Forward and Looking Back: Integrating Completion and Sunk-Cost Effect Within an Escalation-of-Commitment Progress Decision, in: *Journal of Applied Psychology*, 85. Jg., Nr. 1, S. 104–113.
- Moon**, Henry (2001b), The Two Faces of Conscientiousness: Duty and Achievement Striving in Escalation of Commitment Dilemmas, in: *Journal of Applied Psychology*, 86. Jg., S. 535–540.

- Morgan**, Gareth (1988), Accounting as Reality Construction: Towards a New Epistemology for Accounting Practice, in: Accounting, Organizations and Society, 13. Jg., Nr. 5, S. 477–485.
- Muthén**, Bengt, David **Kaplan** und Michael **Hollis** (1987), On Structural Equation Modeling With Data That is Not Missing Completely at Random, in: Psychometrika, 52. Jg., S. 431–462.
- Nathanson**, Sinaia, Joel **Brockner**, Dan **Brenner**, Charles **Samuelson**, Martin **Countryman**, Mary **Lloyd** und Jeffrey Z. **Rubin** (1982), Toward the Reduction of Entrapment, in: Journal of Applied Social Psychology, 12. Jg., Nr. 3, S. 193–208.
- Newell**, Allen und Herbert A. **Simon** (1972), Human Problem Solving, Englewood Cliffs, NJ: Prentice-Hall.
- Newman**, Michael und Rajiv **Sabherwal** (1996), Determinants of Commitment to Information Systems Development: A Longitudinal Investigation, in: Management Information Systems Quarterly, 20. Jg., Nr. 1, S. 23–54.
- Niyonsenga**, Théophile (1997), Response Probability Estimation, in: Journal of Statistical Planning and Inference, 59. Jg., S. 111–126.
- Northcraft**, Gregory B. und Margaret A. **Neale** (1986), Opportunity Costs and the Framing of Resource Allocation Decisions, in: Organizational Behavior and Human Decision Processes, 37. Jg., S. 348–356.
- Northcraft**, Georgy B. und Gerrit **Wolf** (1984), Dollars, Sense, and Sunk Costs: A Life Cycle Model of Resource Allocation Decisions, in: Academy of Management Review, 9. Jg., Nr. 2, S. 225–234.
- Nunnally**, Jum C. (1967), Psychometric Theory, New York, NY: McGraw-Hill.
- Nunnally**, Jum C. (1978), Psychometric Theory, 2. Aufl., New York, NY: McGraw-Hill.
- Nunnally**, Jum C. und Ira H. **Bernstein** (1994), Psychometric Theory, 3. Aufl., New York, NY: McGraw-Hill.
- Nutt**, Paul C. (2002), Why Decisions Fail: Avoiding the Blunders and Traps that Lead to Debacles, San Francisco: Berret-Koehler-Publishers.
- Oskamp**, Stuart (1965), Overconfidence in Case Study Judgments, in: Journal of Consulting Psychology, 29. Jg., S. 261–265.
- Ovid Technologies** (2007), Health and Psychosocial Instruments (HaPI), URL: <http://www.ovid.com/site/catalog/DataBase/866.jsp>; download: April 3, 2007.
- Pallant**, Julie (2005), SPSS Survival Manual, 2. Aufl., Maidenhead: Open University Press.
- Paulhus**, Delroy L. (1984), Two-Component Models of Socially Desirable Responding, in: Journal of Personality and Social Psychology, 46. Jg., S. 598–609.
- Paulhus**, Delroy L. (1986), Self-Deception and Impression Management in Test Responses, in: A. Angleitner und J. S. Wiggins (Hrsg.), Personality Assessment via Questionnaire, New York, NY: Springer, S. 142–165.

- Paulhus**, Delroy L. (1992), Assessing Self-Deception and Impression Management in Self-Reports: The Balanced Inventory of Desirable Responding (Reference Manual, Version 6), Vancouver, BC: University of British Columbia.
- Ping**, Robert A. Jr. (2004), On Assuring Valid Measures for Theoretical Models Using Survey Data, in: *Journal of Business Research*, 57. Jg., S. 125–141.
- Pinto**, Jeffrey K. und Dennis P. **Slevin** (1988), Project Success: Definitions and Measurement Techniques, in: *Project Management Journal*, 19. Jg., Nr. 1, S. 67–75.
- Podsakoff**, Philip M., Scott B. **MacKenzie**, Jeong-Yeon **Lee** und Nathan P. **Podsakoff** (2003), Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies, in: *Journal of Applied Psychology*, 88. Jg., Nr. 5, S. 879–903.
- Popper**, Karl R. (1985), The Rationality Principle, in: D. Miller (Hrsg.), *Popper Selections*, Princeton, NJ: Princeton University Press, S. 357–365.
- Popper**, Karl R. (1994), *The Myth of the Framework: In Defense of Science and Rationality*, London: Routledge.
- Pott**, Philip (1992), *Verlusteskalation und Entscheidungsbindung*, Wiesbaden: DUV.
- Project Management Institute** (Hrsg.) (2006), *A Guide to the Project Management Body of Knowledge*, 3. Aufl., Newtown Square, Pennsylvania: Project Management Institute, Inc.
- Rapoport**, Amnon (1995), Individual Strategies in a Market Entry Game, in: *Group Decision and Negotiation*, 4. Jg., Nr. 2, S. 117–133.
- Rapoport**, Amnon, Darryl A. **Seale**, Ido **Erev** und James A. **Sundali** (1998), Equilibrium Play in Large Group Market Entry Games, in: *Management Science*, 44. Jg., Nr. 1, S. 119–141.
- Raykov**, Tenko und Keith F. **Widaman** (1995), Issues in Applied Structural Equation Modeling Research, in: *Structural Equation Modeling*, 2. Jg., S. 289–318.
- Rieden**, Charlotte (1959), *Eine Analyse der Erinnerungsinhalte in ihrem zeitlichen Verlauf*, Göttingen: Georg-August-Universität Göttingen.
- Riedl**, Josef E. (1990), *Projekt-Controlling in Forschung und Entwicklung*, Berlin: Springer.
- Robbins**, Lionel (1935), *An Essay on the Nature and Significance of Economic Science*, London: MacMillan and Co.
- Rogelberg**, Steven G. und Alexandra **Luong** (1998), Nonresponse to Mailed Surveys: A Review and Guide, in: *Current Directions in Psychological Science*, 7. Jg., S. 60–65.
- Rosenthal**, Robert (1976), *Experimenter Effects in Behavioral Research*, erweiterte Aufl., The Century Psychology Series, New York, NY: Irvington Publishers, Inc.
- Ross**, Jerry und Barry M. **Staw** (1993), Organizational Escalation and Exit: Lessons from the Shoreham Nuclear Power Plant, in: *Academy of Management Journal*, 36. Jg., Nr. 4, August, S. 701–732.
- Rossiter**, John R. (2002), The C-OAR-SE Procedure for Scale Development in Marketing, in: *International Journal of Research in Marketing*, 19. Jg., S. 305–335.

- Rubin**, Donald B. (1976), Inference and Missing Data, in: *Biometrika*, 63. Jg., Nr. 3, S. 581–592.
- Rubin**, Jeffrey Z. und Joel **Brockner** (1975), Factors Affecting Entrapment in Waiting Situations: The Rosencrantz and Guildenstern Effect, in: *Journal of Personality and Social Psychology*, 31. Jg., Nr. 6, S. 1054–103.
- Rubin**, Jeffrey Z., Joel **Brockner**, Susan **Small-Weil** und Sinaia **Nathanson** (1980), Factors Affecting Entry into Psychological Traps, in: *Journal of Conflict Resolution*, 24. Jg., Nr. 3, S. 405–426.
- Rudski**, Jeffrey M. (2002), Hindsight and Confirmation Biases in an Exercise in Telepathy, in: *Psychological Reports*, 91. Jg., Nr. 3, S. 899–906.
- Russo**, J. Edward und Paul J. H. **Schoemaker** (1992), Managing Overconfidence, in: *Sloan Management Review*, 33. Jg., Nr. 2, S. 7–17.
- Salter**, Stephen B. und David J. **Sharp** (2001), Agency Effects and Escalation of Commitment: Do Small National Culture Differences Matter?, in: *International Journal of Accounting*, 36. Jg., S. 33–45.
- Saraph**, Jayant V., P. George **Benson** und Roger G. **Schroeder** (1989), An Instrument for Measuring the Critical Factors of Quality Management, in: *Decision Sciences*, 20. Jg., Nr. 4, S. 810–829.
- Schafer**, Joseph L. und John W. **Graham** (2002), Missing Data: Our View of the State of the Art, in: *Psychological Methods*, 7. Jg., S. 147–177.
- Schermelleh-Engel**, Karin, Helfried **Moosbrugger** und Hans **Müller** (2003), Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures, in: *Methods of Psychological Research Online*, 8. Jg., Nr. 2, S. 23–74.
- Schermelleh-Engel**, Karin und Karl **Schweizer** (2007), Multitrait-Multimethod-Analysen, in: H. Moosbrugger und A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*, Heidelberg: Springer, Kapitel 14, S. 325–341.
- Schermelleh-Engel**, Karin und Christina **Werner** (2007), Methoden der Reliabilitätsbestimmung, in: H. Moosbrugger und A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*, Heidelberg: Springer, Kapitel 6, S. 111–133.
- Schäffer**, Utz (2008), *Management Accounting & Control Scales Handbook*, Wiesbaden: DUV.
- Schmeisser**, Wilhelm, Alexander **Kantner**, Andrea **Geburtig** und Falko **Schindler** (2006), *Forschungs- und Technologie-Controlling: Wie Unternehmen Innovationen operativ und strategisch steuern*, Stuttgart: Schäffer-Poeschel.
- Schmidt**, Jeffrey B. und Roger J. **Calantone** (2002), Escalation of Commitment During New Product Development, in: *Journal of the Academy of Marketing Science*, 30. Jg., Nr. 2, S. 103–118.
- Scholderer**, Joachim und Ingo **Balderjahn** (2006), Was unterscheidet harte und weiche Strukturgleichungsmodelle nun wirklich? Ein Klärungsversuch zur LISREL-PLS-Frage, in: *Marketing – Zeitschrift für Forschung und Praxis*, 28. Jg., Nr. 1, S. 57–70.

- Scholderer**, Joachim, Ingo **Balderjahn** und Marcel **Paulssen** (2006), Kausalität, Linearität, Reliabilität: Drei Dinge, die Sie nie über Strukturgleichungsmodelle wissen wollten, in: *Die Betriebswirtschaft*, 66. Jg., S. 640–650.
- Schoorman**, F. David (1988), Escalation Bias in Performance Appraisals: An Unintended Consequence of Supervisor Participation in Hiring Decisions, in: *Journal of Applied Psychology*, 73. Jg., Nr. 1, S. 58–62.
- Schriesheim**, Chester A. und Kenneth D. **Hill** (1981), Controlling Acquiescence Response Bias by Item Reversal: The Effect on Questionnaire Validity, in: *Educational and Psychological Measurement*, 41. Jg., Nr. 4, S. 1101–1114.
- Schriesheim**, Chester A. und Timothy R. **Hinkin** (1990), Influence Tactics Used by Subordinates: A Theoretical and Empirical Analysis and Refinement of the Kipnis, Schmidt, and Wilkinson Subscales, in: *Journal of Applied Psychology*, 75. Jg., S. 246–257.
- Schriesheim**, Chester A., Angelo J. **Kinicki** und Janet F. **Schriesheim** (1979), The Effect of Leniency on Leader Behavior Descriptions, in: *Organizational Behavior and Human Performance*, 23. Jg., S. 1–29.
- Schriesheim**, Chester A., Kathleen J. **Powers**, Terri A. **Scandura**, Claudia C. **Gardiner** und Melenie J. **Lankau** (1993), Improving Construct Measurement in Management Research: Comments and a Quantitative Approach for Assessing the Theoretical Content Adequacy of Paper-and-Pencil Survey-type Instruments, in: *Journal of Management*, 19. Jg., S. 385–417.
- Schwab**, Donald P. (1980), Construct Validity in Organizational Behavior, in: B. M. Staw und L. L. Cummings (Hrsg.), *Research in Organizational Behavior*, Bd. 2, Greenwich, CT: JAI Press, S. 2–43.
- Schwarz**, Norbert (1999), Self-Reports: How Questions Shape the Answers, in: *American Psychologist*, 54. Jg., Nr. 2, S. 93–105.
- Schweiger**, David M., William R. **Sandberg** und James W. **Ragan** (1986), Group Approaches for Improving Strategic Decision Making: A Comparative Analysis of Dialectical Inquiry, Devil's Advocacy, and Consensus, in: *Academy of Management Journal*, 29. Jg., Nr. 1, S. 51–71.
- Schwenk**, Charles R. (1988), Effects of Devil's Advocacy on Escalating, in: *Human Relations*, 41. Jg., Nr. 10, S. 769–782.
- Sharp**, David J. und Stephen B. **Salter** (1997), Project Escalation and Sunk Costs: A Test of the International Generalizability of Agency and Prospect Theories, in: *Journal of International Business Studies*, 28. Jg., Nr. 1, S. 101–121.
- Sherwood**, Gregory G. (1981), Self-serving Biases in Person Perception: A Reexamination of Projection as a Mechanism of Defense, in: *Psychological Bulletin*, 90. Jg., S. 445–459.
- Simon**, Herbert A. (1957), *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*, New York, NY: Wiley.
- Simon**, Herbert A. (1981), *Entscheidungsverhalten in Organisationen*, Landsberg am Lech: Verlag Moderne Industrie.

- Simon**, Herbert A., George **Kozmetsky**, Harold **Guetzkow** und Gordon **Tyndall** (1954), Centralization vs. Decentralization in Reorganizing the Controller's Department, New York, NY: Controllershship Foundation.
- Simon**, Julian und Rita **Simon** (1975), The Effect of Money Incentives on Family Size: A Hypothetical-Question Study, in: *Public Opinion Quarterly*, 38. Jg., Winter, S. 585–595.
- Simonson**, Itamar und Barry M. **Staw** (1992), Deescalation Strategies: A Comparison of Techniques for Reducing Commitment to Losing Courses of Action, in: *Journal of Applied Psychology*, 77. Jg., Nr. 4, S. 419–426.
- Singer**, Wolf (2002a), Der Beobachter im Gehirn, in: *Der Beobachter im Gehirn: Essays zur Hirnforschung*, Frankfurt am Main: Suhrkamp, S. 144–170.
- Singer**, Wolf (2002b), Wahrnehmen, Erinnern, Vergessen, in: *Der Beobachter im Gehirn: Essays zur Hirnforschung*, Frankfurt am Main: Suhrkamp, S. 77–86.
- Smith**, H. Jeff und Mark **Keil** (2003), The Reluctance to Report Bad News on Troubled Software Projects: A Theoretical Model, in: *Information Systems Journal*, 13. Jg., S. 69–95.
- Smith**, H. Jeff, Mark **Keil** und Gordon **Depledge** (2001), Keeping Mum as the Project Goes Under: Toward an Explanatory Model, in: *Journal of Management Information Systems*, 18. Jg., Nr. 2, S. 189–227.
- Smith**, Vernon L. (1979), An Experimental Comparison of Three Public Good Decision Mechanisms, in: *Scandinavian Journal of Economics*, 81. Jg., S. 198–215.
- Sorter**, George H., Selwyn W. **Becker**, T. Ross **Archibald** und William H. **Beaver** (1966), Accounting and Financial Measures as Indicators, in: R. K. Jaedicke, Y. Ijiri und O. Nielsen (Hrsg.), *Research in Accounting Measurement*, Sarasota, FL: American Accounting Association, S. 200–210.
- Spector**, Paul E. (1992), Summated Rating Scale Construction: An Introduction, Bd. 07-082 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Newbury Park, CA: Sage Publications.
- Spector**, Paul E. (2006), Method Variance in Organizational Research: Truth or Urban Legend, in: *Organizational Research Method*, 9. Jg., Nr. 2, S. 221–232.
- Staw**, Barry M. (1976), Knee-Deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action, in: *Organizational Behavior and Human Performance*, 16. Jg., S. 27–44.
- Staw**, Barry M. (1981), The Escalation of Commitment To a Course of Action, in: *Academy of Management Review*, 6. Jg., Nr. 4, S. 577–587.
- Staw**, Barry M. (1997), Escalation of Commitment: An Update and Appraisal, in: Z. Shapira (Hrsg.), *Organizational Decision Making*, Cambridge, UK: Press Syndicate of the University of Cambridge, S. 191–215.
- Staw**, Barry M. und Jerry **Ross** (1987a), Behavior in Escalation Situations: Antecedents, Prototypes, and Solutions, in: L. L. Cummings und B. M. Staw (Hrsg.), *Research in Organization Behavior*, Greenwich, CT: JAI Press, S. 39–78.

- Staw**, Barry M. und Jerry **Ross (1987b)**, Knowing When to Pull the Plug, in: Harvard Business Review, 65. Jg., S. 68–74.
- Stedry, Andrew, C. (1960)**, Budget Control and Cost Behavior, The Ford Foundation Doctoral Dissertation Series, Englewood Cliffs, NJ: Prentice-Hall.
- Steeh**, Charlotte G. (1981), Trends in Nonresponse Rates, 1952-1979, in: Public Opinion Quarterly, 45. Jg., S. 40–57.
- Steiger**, James H. (1990), Structural Model Evaluation and Modification: An Interval Estimation Approach, in: Multivariate Behavioral Research, 25. Jg., S. 173–180.
- Steiger**, James H. und John C. **Lind (1980)**, Statistically Based Tests for the Number of Common Factors, Paper Presented at the Annual Spring Meeting of the Psychometric Society, Iowa City.
- Strickland**, Lloyd, Roy J. **Lewici** und Arnold M. **Katz (1966)**, Temporal Orientation and Perceived Control as Determinants of Risk Taking, in: Journal of Experimental Social Psychology, 2. Jg., S. 143–151.
- Svenson**, Ola (1981), Are We All Less Risky and More Skillful Than Our Fellow Drivers, in: Acta Psychologica, 47. Jg., Nr. 2, S. 143–148.
- Swieringa**, Robert J. und Karl E. **Weik (1982)**, An Assessment of Laboratory Experiments in Accounting, in: Journal of Accounting Research, Supplement 1982, 20. Jg., S. 56–101.
- Tabachnick**, Barbara G. und Linda **Fidell (2001)**, Using Multivariate Statistics, 4. Aufl., New York, NY: HarperCollins.
- Tadisina**, Suresh K. (1986), Support System for the Determination Decision in R&D Management, in: Project Management Journal, 17. Jg., S. 97–104.
- Tan**, Bernhard C. Y., H. Jeff **Smith**, Mark **Keil** und Ramiro **Montealegre (2003)**, Reporting Bad News About Software Projects: Impact of Organizational Climate and Information Asymmetry in an Individualistic and a Collectivistic Culture, in: IEEE Transactions on Engineering Management, 50. Jg., Nr. 1, Februar, S. 64–77.
- Teger**, Allan I. (1980), Too Much Invested To Quit, New York, NY: Pergamon Press.
- Temme**, Dirk (2006), Die Spezifikation und Identifikation formativer Messmodelle der Marketingforschung in Kovarianzstrukturanalysen, in: Marketing – Zeitschrift für Forschung und Praxis, 28. Jg., Nr. 3, S. 183–196.
- Thompson**, James D. (1967), Organizations in Action, New York, NY: McGraw-Hill.
- Tiwana**, Amrit (2000), The Knowledge Management Toolkit: Techniques for Building a Knowledge Management System, Upper Saddle River, NJ: Prentice Hall.
- Tiwana**, Amrit, Mark **Keil** und Robert G. **Fichman (2006)**, Information Systems Project Continuation in Escalation Situations: A Real Options Model, in: Decision Sciences, 37. Jg., Nr. 3, August, S. 357–391.
- Tomassini**, L. A., Ira **Solomon**, J. L. **Krogstad** und M. B. **Romney (1982)**, Calibration of Auditors' Probabilistic Judgments: Some Empirical Evidence, in: Organizational Behavior and Human Performance, 30. Jg., S. 391–406.

- Tucker**, Ledyard R. und Charles **Lewis** (1973), A Reliability Coefficient for Maximum Likelihood Factor Analysis, in: *Psychometrika*, 38. Jg., Nr. 3, S. 1–10.
- Tversky**, Amos und Daniel **Kahneman** (1974), Judgement Under Uncertainty: Heuristics and Biases, in: *Science*, 185. Jg., S. 1124–1131.
- Vandenbosch**, Mark B. (1996), Confirmatory Compositional Approaches to the Development of Product Spaces, in: *European Journal of Marketing*, 30. Jg., Nr. 3, S. 23–46.
- Wallace**, Linda und Mark **Keil** (2004), Software Project Risks and Their Effect on Outcomes, in: *Communications of the Association for Computing Machinery*, 47. Jg., Nr. 4, S. 68–73.
- Wallace**, Linda, Mark **Keil** und Arun **Rai** (2004a), How Software Project Risk Affects Project Performance: An Investigation of the Dimensions of Risk and an Exploratory Model, in: *Decision Sciences*, 35. Jg., Nr. 2, S. 289–321.
- Wallace**, Linda, Mark **Keil** und Arun **Rai** (2004b), Understanding Software Project Risk: A Cluster Analysis, in: *Information & Management*, 42. Jg., Nr. 1, S. 115–125.
- Walsh**, James P. (1988), Selectivity and Selective Perception: An Investigation of Managers' Belief Structures and Information Processing, in: *Academy of Management Journal*, 31. Jg., Nr. 4, S. 873–896.
- Watson**, Hugh J. und Barbara J. **Haley** (1997), Data Warehousing: A Framework and Survey of Practices, in: *Journal of Data Warehousing*, 2. Jg., S. 10–17.
- Weber**, Jürgen (2002), Was machen Controller wann warum?, Vallendar: WHU – Otto Beisheim School of Management.
- Weber**, Jürgen, Bernhard **Hirsch**, Stefan **Linder** und Eric **Zayer** (2003), Verhaltensorientiertes Controlling: Der Mensch im Mittelpunkt, Bd. 34 der Reihe Advanced Controlling, Weinheim: Wiley.
- Weber**, Jürgen, Matthias **Meyer**, Holger **Birl**, Ramon **Knollmann**, Hendrik **Schlüter** und Carsten **Sieber** (2006), Investitionscontrolling in deutschen Großunternehmen, Bd. 52 der Reihe Advanced Controlling, Weinheim: Wiley.
- Weber**, Jürgen und Utz **Schäffer** (2006), Einführung in das Controlling, 11. Aufl., Stuttgart: Schäffer-Poeschel.
- Weinstein**, Neil D. (1980), Unrealistic Optimism About Future Life Events, in: *Journal of Personality and Social Psychology*, 39. Jg., S. 806–820.
- Weinstein**, Neil D. (1987), Unrealistic Optimism about Susceptibility to Health Problems, in: *Journal of Behavioral Medicine*, 10. Jg., S. 481–500.
- Weinstein**, Neil D. (1996), Unrealistic Optimism: Present and Future, in: *Journal of Social and Clinical Psychology*, 15. Jg., S. 1–8.
- West**, Stephen G., John F. **Finch** und Patrick J. **Curran** (1995), Structural Equation Models With Nonnormal Variables: Problems and Remedies, in: R. H. Hoyle (Hrsg.), *Structural Equation Modeling: Concepts, Issues and Applications*, Thousand Oaks, CA: Sage, Kapitel 4, S. 56–75.

- Whyte, Glen (1986)**, Escalating Commitment to a Course of Action: A Reinterpretation, in: *Academy of Management Review*, 11. Jg., Nr. 2, S. 311–321.
- Whyte, William H. (1956)**, *The Organization Man*, New York, NY: Simon & Schuster.
- Wildt, Albert R. und Olli Ahtola (1978)**, *Analysis of Covariance*, Bd. 07-012 der Reihe Sage University Paper Series on Quantitative Applications in the Social Sciences, Newbury Park, CA: Sage.
- Wong, Kin F. E. (2005)**, The Role of Risk in Making Decisions Under Escalation Situations, in: *Journal of Applied Psychology*, 54. Jg., Nr. 4, S. 584–607.
- Wouters, Marc, James C. Anderson und Finn Wynstra (2005)**, The Adoption of Total Cost of Ownership for Sourcing Decisions: A Structural Equation Analysis, in: *Accounting, Organizations and Society*, 30. Jg., S. 167–191.
- Yuan, Ke-Hai, Peter M. Bentler und Wai Chan (2004)**, Structural Equation Modeling With Heavy Tailed Distributions, in: *Psychometrika*, 69. Jg., S. 421–436.
- Zayer, Eric C. (2007)**, *Verspätete Projektabbrüche in F&E: Eine verhaltensorientierte Analyse*, Wiesbaden: DUV.
- Zhang, G. Peter, Mark Keil, Arun Rai und Joan Mann (2003)**, Predicting Information Technology Project Escalation: A Neural Network Approach, in: *European Journal of Operational Research*, 146. Jg., S. 115–129.
- Zimbardo, Philip G. und Richard J. Gerig (1996)**, *Psychologie*, Berlin: Springer.